

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-255991

(43)Date of publication of application : 10.09.2003

(51)Int.Cl.

G10L 15/22

A63H 11/00

B25J 5/00

B25J 13/00

G10L 13/00

G10L 15/00

G10L 15/06

G10L 15/20

G10L 17/00

(21)Application number : 2002-060428

(71)Applicant : SONY CORP

(22)Date of filing : 06.03.2002

(72)Inventor : AOYAMA KAZUMI
SHIMOMURA HIDEKI
YAMADA KEIICHI

(54) INTERACTIVE CONTROL SYSTEM, INTERACTIVE CONTROL METHOD, AND ROBOT APPARATUS

(57)Abstract:

PROBLEM TO BE SOLVED: To provide an interactive control system, interactive control method and robot apparatus capable of improving entertainment characteristics.

SOLUTION: The interactive control system formed by connecting a robot and an information processor through the network is arranged to form the history data relating to a play on words among the utterance contents of a user when the interaction by the play on words is made between the robot and the user and to send the data to the information processor. The information processor selectively reads out the content data optimum for the user in accordance with the history data from memory means and provides the original robot with the data.

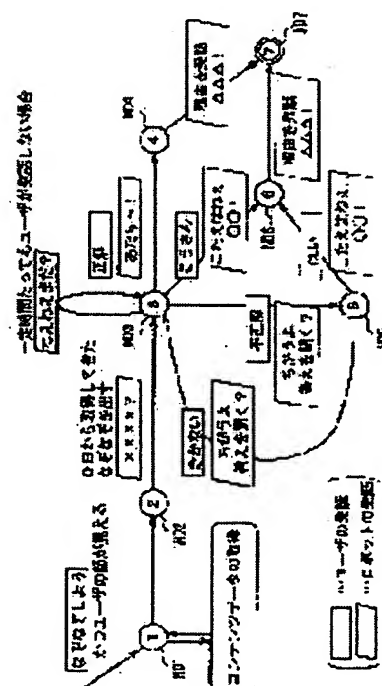


図28 言葉遊びによる対話シーケンス (1)

LEGAL STATUS

[Date of request for examination] 17.03.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than abandonment the examiner's decision of rejection or application converted registration]

[Date of final disposal for application] 29.09.2005

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

【特許請求の範囲】

【請求項1】ロボット及び情報処理装置がネットワークを介して接続された対話制御システムにおいて、上記ロボットに設けられ、人間と対話するための機能を有し、当該対話を通じて対象とするユーザの発話を認識する対話手段と、上記対話手段による上記ユーザの発話内容のうち、言葉遊びに関する履歴データを生成する生成手段と、上記生成手段により生成された上記履歴データを、上記言葉遊びを通じて得られる上記ユーザの発言内容に応じて更新する更新手段と、上記言葉遊びの開始の際には、上記履歴データを上記ネットワークを介して上記情報処理装置に送信する通信手段とを具え、上記情報処理装置に設けられ、複数の上記言葉遊びの内容を表す内容データを記憶する記憶手段と、上記通信手段を介して送信された上記履歴データを検出する検出手段と、上記検出手段によって検出された上記履歴データに基づいて、上記記憶手段から上記内容データを選択的に読み出して上記ネットワークを介して元の上記ロボットに送信する通信制御手段とを具え、上記ロボットの上記対話手段は、上記情報処理装置の上記通信制御手段から送信された上記内容データに基づく上記言葉遊びの内容を出力するを具えることを特徴とする対話制御システム。

【請求項2】上記ロボットでは、上記対話手段は、上記ユーザに出力した上記内容データに基づく上記言葉遊びの内容に関する評価を当該ユーザの発話から認識し、上記更新手段は上記履歴データを上記評価に応じて更新し、上記通信手段は、上記更新手段によって更新された上記履歴データを上記情報処理装置に送信し、上記情報処理装置では、上記記憶手段は、上記言葉遊びの内容データに付随する付随データを当該内容データに関連付けて記憶しておき、上記通信制御手段は、上記選択された上記内容データに付随する上記付随データについて、上記通信手段から送信された上記履歴データに基づく評価に関連するデータ部分を更新することを特徴とする請求項1に記載の対話制御システム。

【請求項3】上記ロボットでは、上記対話手段は、上記ユーザに出力した新たな言葉遊びの内容を当該ユーザの発話から認識し、上記通信手段は、上記言葉遊びの内容を表す新規内容データを上記情報処理装置に送信し、上記情報処理装置では、上記記憶手段は、対応する上記ユーザについての上記内

容データに追加して、上記通信手段から送信された上記新規内容データを記憶することを特徴とする請求項1に記載の対話制御システム。

【請求項4】上記記憶手段は、複数の上記ロボットが共有できるようになされたデータベースであることを特徴とする請求項1に記載の対話制御システム。

【請求項5】ロボット及び情報処理装置がネットワークを介して接続された対話制御方法において、上記ロボットでは、人間との対話を通じて対象とするユーザの発話を認識し、当該ユーザの発話内容のうち、言葉遊びに関する履歴データを生成し、当該生成された履歴データを、上記言葉遊びを通じて得られる上記ユーザの発言内容に応じて更新しながら、上記言葉遊びの開始の際には上記ネットワークを介して上記情報処理装置に送信する第1のステップと、

上記情報処理装置では、予め記憶された複数の上記言葉遊びの内容を表す内容データのうち、上記ロボットから送信された上記履歴データに基づいて選択した上記内容データを読み出して、上記ネットワークを介して元の上記ロボットに送信する第2のステップと、上記ロボットでは、上記情報処理装置から送信された上記内容データに基づく上記言葉遊びの内容を出力する第3のステップとを具えることを特徴とする対話制御方法。

【請求項6】上記第1のステップでは、上記ユーザに出力した上記内容データに基づく上記言葉遊びの内容に関する評価を当該ユーザの発話から認識した後、当該評価に応じて上記履歴データを更新し、当該更新された履歴データを上記情報処理装置に送信し、上記第2のステップでは、上記言葉遊びの内容データに付随する付随データを当該内容データに関連付けて記憶しておき、上記選択された上記内容データに付随する上記付随データについて、上記送信された上記履歴データに基づく評価に関連するデータ部分を更新することを特徴とする請求項5に記載の対話制御方法。

【請求項7】上記第1のステップでは、上記ユーザに出力した新たな言葉遊びの内容を当該ユーザの発話から認識した後、上記言葉遊びの内容を表す新規内容データを上記情報処理装置に送信し、上記第2のステップでは、対応する上記ユーザについての上記内容データに追加して、上記通信手段から送信された上記新規内容データを記憶することを特徴とする請求項5に記載の対話制御方法。

【請求項8】上記第2のステップでは、予め記憶された複数の上記言葉遊びの内容を表す内容データを、複数の上記ロボットが共有できるようにデータベース管理することを特徴とする請求項5に記載の対話制御方法。

【請求項9】情報処理装置とネットワークを介して接続

されたロボット装置において、人間と対話するための機能を有し、当該対話を通じて対象とするユーザの発話を認識する対話手段と、上記対話手段による上記ユーザの発話内容のうち、言葉遊びに関する履歴データを生成する生成手段と、上記生成手段により生成された上記履歴データを、上記言葉遊びを通じて得られる上記ユーザの発言内容に応じて更新する更新手段と、上記言葉遊びの開始の際には、上記履歴データを上記ネットワークを介して上記情報処理装置に送信する通信手段とを具え、上記情報処理装置において予め記憶された複数の上記言葉遊びの内容を表す内容データのうち、上記通信手段から送信された上記履歴データに基づいて選択された上記内容データが上記ネットワークを介して送信されたとき、上記対話手段は、当該内容データに基づく上記言葉遊びの内容を出力するを具えることを特徴とするロボット装置。

【請求項 10】上記対話手段は、上記ユーザに出力した上記内容データに基づく上記言葉遊びの内容に関する評価を当該ユーザの発話から認識し、上記更新手段は上記履歴データを上記評価に応じて更新し、上記通信手段は、上記更新手段によって更新された上記履歴データを上記情報処理装置に送信し、上記情報処理装置では、予め記憶された上記言葉遊びの内容データに付随しかつ当該内容データに関連付けられた付随データのうち、上記選択された上記内容データに付随する上記付随データについて、上記通信手段から送信された上記履歴データに基づく評価に関連するデータ部分が更新されることを特徴とする請求項 9 に記載のロボット装置。

【請求項 11】上記対話手段は、上記ユーザに出力した新たな言葉遊びの内容を当該ユーザの発話から認識し、上記通信手段は、上記言葉遊びの内容を表す新規内容データを上記情報処理装置に送信し、上記情報処理装置では、対応する上記ユーザについての上記内容データに追加して、上記通信手段から送信された上記新規内容データが記憶されることを特徴とする請求項 9 に記載のロボット装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は対話制御システム、対話制御方法及びロボット装置に関し、例えばエンターテインメントロボットに適用して好適なものである。

【0002】

【従来の技術】近年、一般家庭向けのエンターテインメントロボットが多く企業等において開発され、商品化されている。そしてこのようなエンターテインメントロボットの中には、CCD (Charge Coupled Device) カメラ

やマイクロホン等の各種外部センサが搭載され、これら外部センサの出力に基づいて外部状況を認識し、認識結果に基づいて自律的に行動し得るようになされたものなどもある。

【0003】かかるロボットとユーザとが音声による対話を行う音声対話システムを構築する場合、例えばテレフォンショッピングの受け付けや、電話番号案内など、あるタスクを達成することを目的とした音声対話システムが考えられる。

【0004】

【発明が解決しようとする課題】ところが、ロボットと人間が日常的に会話する場面を想定したとき、ロボットは、タスク達成のための対話のほかに、雑談や言葉遊びなど、毎日会話しても飽きないような会話ができればならないのであるが、上述のようなタスクの遂行を目的とする対話システムでは、システム内の電話番号リストやショッピングアイテムリストなどのデータが特定内容に固定されているため、ロボットの会話に面白みを持たせることができず、さらにはシステムを使用する個人の好みに応じて当該システム内のデータを変更することもできなかった。

【0005】特に、ロボットと人間が日常的な会話として、なぞなぞや山手線ゲーム（特定の事項に関連する内容の言葉を互いに重複しないように順番に言い合う遊び）等の言葉遊びによる対話を行う場合、ユーザを飽きさせないためには、ロボットは大量の対話内容（コンテンツ）を表すデータ（以下、これをコンテンツデータと呼ぶ）を保持する必要がある。

【0006】そこで近年では、インターネット上に分散する各サーバ内の各種情報を相互に関連付けて検索可能にした情報網であるWeb（すなわちWWW: World Wide Web）が、情報サービスとして幅広く利用されており、かかるWebを利用して、大量のコンテンツを保有するコンテンツサーバが、ロボットとの間で当該ロボットが持つべきコンテンツデータのやり取りを行うことにより、当該ロボットと対面するユーザが日常的な会話を行うことができると考えられる。

【0007】かかるコンテンツサーバは、大量のコンテンツデータを利用可能な全てのロボットが共有できるデータベースに格納しており、必要に応じて当該データベースから対応するコンテンツデータを読み出してネットワークを介してロボットに発話させ得るよう構築されている。

【0008】しかし、実際にロボットとユーザとの間で言葉遊びを行う際には、個々のユーザはそれぞれ好みや難易度に対するスキルが多種多様であるため、該当するロボットがデータベースに格納されている大量のコンテンツデータの中からランダムにコンテンツデータを取得する手法では、全てのユーザのニーズに十分に答えられないといった問題があった。

【0009】この問題を解決する一つの方法として、ユーザの好みやレベルを表すプロフィール情報と、コンテンツに付随する内容の分類情報とを、データベースに格納しておき、コンテンツサーバがロボットからの要求に応じてデータベースからユーザが所望するコンテンツデータを取得するときに、プロフィール情報及び分類情報に関連のあるコンテンツデータを選択するようにする方法が考えられる。

【0010】ところが、なぞなぞや山手線ゲーム等の言葉遊びを目的とする対話では、ロボットとユーザとの間に、会話のリズムや面白さといったものが要求されるのであるが、現在の音声認識処理の技術では、ユーザの発話に対する認識間違いを避けることができず、ロボットがいちいちユーザの発話内容を確認的に発するのでは、ユーザとの会話が不自然な状態になってしまうおそれがある。

【0011】例えばロボットが「2回食べると元気になる食べ物なんだ？」というなぞなぞを出題したときに、ユーザが「のり」と答えた場合、ロボットが「答えはのりですね」というように直接的な確認をする旨の発現をしてしまうのは、会話の流れを止めると同時に面白みに欠けてしまう。

【0012】これに対してロボットがユーザの発話内容を無視して会話を続けるのでは、ユーザ自身が自分の発話内容をロボットがどのように認識したのかを確認することできず、会話中に不安感を与えるおそれがあった。

【0013】本発明は以上の点を考慮してなされたもので、エンターテインメント性を格段的に向上させ得る対話制御システム、対話制御方法及びロボット装置を提案しようとするものである。

【0014】

【課題を解決するための手段】かかる課題を解決するため本発明においては、ロボット及び情報処理装置がネットワークを介して接続された対話制御システムにおいて、ロボットには、人間と対話するための機能を有し、当該対話を通じて対象とするユーザの発話を認識する対話手段と、対話手段によるユーザの発話内容のうち、言葉遊びに関する履歴データを生成する生成手段と、生成手段により生成された履歴データを、言葉遊びを通じて得られるユーザの発言内容に応じて更新する更新手段と、言葉遊びの開始の際には、履歴データをネットワークを介して情報処理装置に送信する通信手段とを設け、また情報処理装置には、複数の言葉遊びの内容を表す内容データを記憶する記憶手段と、通信手段を介して送信された履歴データを検出する検出手段と、検出手段によって検出された履歴データに基づいて、記憶手段から内容データを選択的に読み出してネットワークを介して元のロボットに送信する通信制御手段とを設けるようにした。そしてロボットの対話手段は、情報処理装置の通信制御手段から送信された内容データに基づく言葉遊びの

内容を入力するようにした。

【0015】この結果、この対話制御システムでは、ロボットとユーザとの間で言葉遊びによる対話をする際、ユーザの発話内容のうち言葉遊びに関する履歴データを生成して情報処理装置に送信し、当該情報処理装置が記憶手段から当該履歴データに基づいてユーザに最適な内容データを選択的に読み出して元のロボットに提供するようにしたことにより、ユーザとの間でロボットの会話に面白みやリズムを持たせることができ、あたかも人間同士が会話しているかのごとく自然な日常会話に近づけることができる。

【0016】また本発明においては、ロボット及び情報処理装置がネットワークを介して接続された対話制御方法において、ロボットでは、人間との対話を通じて対象とするユーザの発話を認識し、当該ユーザの発話内容のうち、言葉遊びに関する履歴データを生成し、当該生成された履歴データを、言葉遊びを通じて得られるユーザの発言内容に応じて更新しながら、言葉遊びの開始の際にはネットワークを介して情報処理装置に送信する第1のステップと、情報処理装置では、予め記憶された複数の言葉遊びの内容を表す内容データのうち、ロボットから送信された履歴データに基づいて選択した内容データを読み出して、ネットワークを介して元のロボットに送信する第2のステップと、ロボットでは、情報処理装置から送信された内容データに基づく言葉遊びの内容を入力する第3のステップとを設けるようにした。

【0017】この結果、この対話制御方法では、ロボットとユーザとの間で言葉遊びによる対話をする際、ユーザの発話内容のうち言葉遊びに関する履歴データを生成して情報処理装置に送信し、当該情報処理装置が履歴データに基づいてユーザに最適な内容データを複数の内容データの中から選択的に読み出して元のロボットに提供するようにしたことにより、ユーザとの間でロボットの会話に面白みやリズムを持たせることができ、あたかも人間同士が会話しているかのごとく自然な日常会話に近づけることができる。

【0018】さらに本発明においては、情報処理装置とネットワークを介して接続されたロボット装置において、人間と対話するための機能を有し、当該対話を通じて対象とするユーザの発話を認識する対話手段と、対話手段によるユーザの発話内容のうち、言葉遊びに関する履歴データを生成する生成手段と、生成手段により生成された履歴データを、言葉遊びを通じて得られるユーザの発言内容に応じて更新する更新手段と、言葉遊びの開始の際には、履歴データをネットワークを介して情報処理装置に送信する通信手段とを設け、情報処理装置において予め記憶された複数の言葉遊びの内容を表す内容データのうち、通信手段から送信された履歴データに基づいて選択された内容データがネットワークを介して送信されたとき、対話手段は、当該内容データに基づく言葉

遊びの内容を出力するようにした。

【0019】この結果、このロボット装置では、ロボットとユーザとの間で言葉遊びによる対話をする際、ユーザの発話内容のうち言葉遊びに関する履歴データを生成して情報処理装置に送信し、当該情報処理装置から履歴データに基づくユーザに最適な内容データを選択的に取得するようにしたことにより、ユーザとの間でロボットの会話に面白みやリズムを持たせることができ、あたかも人間同士が会話しているかのごとく自然な日常会話に近づけることができる。

【0020】

【発明の実施の形態】以下図面について、本発明の一実施の形態を詳述する。

【0021】(1)本実施の形態によるロボットの構成図1及び図2において、1は全体として本実施の形態による2足歩行型のロボットを示し、胴体部ユニット2の上部に頭部ユニット3が配設されると共に、当該胴体部ユニット2の上部左右にそれぞれ同じ構成の腕部ユニット4A、4Bがそれぞれ配設され、かつ胴体部ユニット2の下部左右にそれぞれ同じ構成の脚部ユニット5A、5Bがそれぞれ所定位置に取り付けられることにより構成されている。

【0022】胴体部ユニット2においては、体幹上部を形成するフレーム10及び体幹下部を形成する腰ベース11が腰関節機構12を介して連結することにより構成されており、体幹下部の腰ベース11に固定された腰関節機構12の各アクチュエータA1、A2をそれぞれ駆動することによって、体幹上部を図3に示す直交するロール軸13及びピッチ軸14の回りにそれぞれ独立に回転させることができるようになされている。

【0023】また頭部ユニット3は、フレーム10の上端に固定された肩ベース15の上面中央部に首関節機構16を介して取り付けられており、当該首関節機構16の各アクチュエータA3、A4をそれぞれ駆動することによって、図3に示す直交するピッチ軸17及びヨー軸18の回りにそれぞれ独立に回転させることができるようになされている。

【0024】さらに各腕部ユニット4A、4Bは、それぞれ肩関節機構19を介して肩ベース15の左右に取り付けられており、対応する肩関節機構19の各アクチュエータA5、A6をそれぞれ駆動することによって図3に示す直交するピッチ軸20及びロール軸21の回りにそれぞれ独立に回転させることができるようになされている。

【0025】この場合、各腕部ユニット4A、4Bは、それぞれ上腕部を形成するアクチュエータA7の出力軸に肘関節機構22を介して前腕部を形成するアクチュエータA8が連結され、当該前腕部の先端に手部23が取り付けられることにより構成されている。

【0026】そして各腕部ユニット4A、4Bでは、ア

クチュエータA7を駆動することによって前腕部を図3に示すヨー軸24の回りに回転させ、アクチュエータA8を駆動することによって前腕部を図3に示すピッチ軸25の回りにそれぞれ回転させることができるようになされている。

【0027】これに対して各脚部ユニット5A、5Bにおいては、それぞれ股関節機構26を介して体幹下部の腰ベース11にそれぞれ取り付けられており、それぞれ対応する股関節機構26の各アクチュエータをA9～A11それぞれ駆動することによって、図3に示す互いに直交するヨー軸27、ロール軸28及びピッチ軸29の回りにそれぞれ独立に回転させることができるようになされている。

【0028】この場合各脚部ユニット5A、5Bは、それぞれ大腿部を形成するフレーム30の下端に膝関節機構31を介して下腿部を形成するフレーム32が連結されると共に、当該フレーム32の下端に足首関節機構33を介して足部34が連結されることにより構成されている。

【0029】これにより各脚部ユニット5A、5Bにおいては、膝関節機構31を形成するアクチュエータA12を駆動することによって、下腿部を図3に示すピッチ軸35の回りに回転させることができ、また足首関節機構33のアクチュエータA13、A14をそれぞれ駆動することによって、足部34を図3に示す直交するピッチ軸36及びロール軸37の回りにそれぞれ独立に回転させることができるようになされている。

【0030】一方、胴体部ユニット2の体幹下部を形成する腰ベース11の背面側には、図4に示すように、当該ロボット1全体の動作制御を司るメイン制御部40と、電源回路及び通信回路などの周辺回路41と、バッテリー45(図5)となどがボックスに収納されてなる制御ユニット42が配設されている。

【0031】そしてこの制御ユニット42は、各構成ユニット(胴体部ユニット2、頭部ユニット3、各腕部ユニット4A、4B及び各脚部ユニット5A、5B)内にそれぞれ配設された各サブ制御部43A～43Dと接続されており、これらサブ制御部43A～43Dに対して必要な電源電圧を供給したり、これらサブ制御部43A～43Dと通信を行ったりすることができるようになされている。

【0032】また各サブ制御部43A～43Dは、それぞれ対応する構成ユニット内の各アクチュエータA1～A14と接続されており、当該構成ユニット内の各アクチュエータA1～A14をメイン制御部40から与えられる各種制御コマンドに基づいて指定された状態に駆動し得るようになされている。

【0033】さらに頭部ユニット3には、図5に示すように、このロボット1の「目」として機能するCCD(Charge Coupled Device)カメラ50及び「耳」とし

て機能するマイクロホン51及びタッチセンサ52などからなる外部センサ部53と、「口」として機能するスピーカ54などがそれぞれ所定位置に配設され、制御ユニット42内には、バッテリーセンサ55及び加速度センサ56などからなる内部センサ部57が配設されている。

【0034】そして外部センサ部53のCCDカメラ50は、周囲の状況を撮像し、得られた画像信号S1Aをメイン制御部に送出する一方、マイクロホン51は、ユーザから音声入力として与えられる「歩け」、「伏せ」又は「ボールを追いかけて」等の各種命令音声を集音し、かくして得られた音声信号S1Bをメイン制御部40に送出するようになされている。

【0035】またタッチセンサ52は、図1及び図2において明らかなように頭部ユニット3の上部に設けられており、ユーザからの「撫でる」や「叩く」といった物理的な働きかけにより受けた圧力を検出し、検出結果を圧力検出信号S1Cとしてメイン制御部40に送出する。

【0036】さらに内部センサ部57のバッテリーセンサ55は、バッテリー45のエネルギー残量を所定周期で検出し、検出結果をバッテリー残量検出信号S2Aとしてメイン制御部40に送出する一方、加速度センサ56は、3軸方向(x軸、y軸及びz軸)の加速度を所定周期で検出し、検出結果を加速度検出信号S2Bとしてメイン制御部40に送出する。

【0037】メイン制御部40は、外部センサ部53のCCDカメラ50、マイクロホン51及びタッチセンサ52等からそれぞれ供給される画像信号S1A、音声信号S1B及び圧力検出信号S1C等(以下、これらをまとめて外部センサ信号S1と呼ぶ)と、内部センサ部57のバッテリーセンサ55及び加速度センサ等からそれぞれ供給されるバッテリー残量検出信号S2A及び加速度検出信号S2B等(以下、これらをまとめて内部センサ信号S2と呼ぶ)に基づいて、ロボット1の周囲及び内部の状況や、ユーザからの指令、ユーザからの働きかけの有無などを判断する。

【0038】そしてメイン制御部40は、この判断結果と、予め内部メモリ40Aに格納されている制御プログラムと、そのとき装填されている外部メモリ58に格納されている各種制御パラメータとに基づいて続く行動を決定し、決定結果に基づく制御コマンドを対応するサブ制御部43A~43Dに送出する。この結果、この制御コマンドに基づき、そのサブ制御部43A~43Dの制御のもとに、対応するアクチュエータA1~A14が駆動され、かくして頭部ユニット3を上下左右に揺動させたり、腕部ユニット4A、4Bを上にあげたり、歩行するなどの行動がロボット1により発現されることとなる。

【0039】またこの際メイン制御部40は、必要に応

じて所定の音声信号S3をスピーカ54に与えることにより当該音声信号S3に基づく音声を外部に出力させたり、外見上の「目」として機能する頭部ユニット3の所定位置に設けられたLEDに駆動信号を出力することによりこれを点滅させる。

【0040】このようにしてこのロボット1においては、周囲及び内部の状況や、ユーザからの指令及び働きかけの有無などに基づいて自律的に行動することができるようになされている。

【0041】(2) 本実施の形態による対話制御システムの構成

ここで図6は、ユーザが所有する複数の上述したロボット1と、情報提供側60が配置したコンテンツサーバ61とがネットワーク62を介して接続されることにより構成される本実施の形態による対話制御システム63を示すものである。

【0042】各ロボット1においては、ユーザからの指令や周囲の環境に応じて自律的に行動する一方、コンテンツサーバ61とネットワーク62を介して通信することにより必要なデータを送受信したり、当該通信により得られたコンテンツデータ等に基づく音声をスピーカ54(図5)を介して放音することができるようになされている。

【0043】実際に各ロボット1には、例えばCD(Compact Disc)-ROMに記録されて提供される、この対話制御システム63全体としてかかる機能を発揮させるためのアプリケーションソフトウェアがインストールされると共に、例えばブルートゥース(Bluetooth)等の所定の無線通信規格に対応した無線LANカード(図示せず)が胴体部ユニット2(図1)内の所定部位に装着されるようになされている。

【0044】またコンテンツサーバ61は、情報提供側60が提供する後述のような各種サービスに関する各種処理を行うWebサーバ及びデータベースサーバであり、ネットワーク62を介してアクセスしてきたロボット1と通信して必要なデータを送受信することができるようになされている。

【0045】なおコンテンツサーバ61の構成を図7に示す。この図7からも明らかなように、コンテンツサーバ61は、コンテンツサーバ61全体の制御を司るCPU65と、各種ソフトウェアが格納されたROM66と、CPU65のワークメモリとしてのRAM67と、各種データが格納されたハードディスク装置68と、CPU65がネットワーク62(図6)を介して外部と通信するためのインターフェースであるネットワークインターフェース部69とを有し、これらがバス70を介して相互に接続されることにより構成されている。

【0046】この場合CPU65は、ネットワーク62を介してアクセスしてきたロボット1から与えられるデータやコマンドをネットワークインターフェース部69

を介して取り込み、当該データやコマンドと、ROM 66に格納されているソフトウェアとに基づいて各種処理を実行する。このネットワークインターフェース部69は、例えばブルートゥース(Bluetooth)等の無線LAN方式で各種データをやり取りするLAN制御部(図示せず)を有する。

【0047】そしてCPU 65は、この処理結果として、例えばハードディスク装置68から読み出した所定のWebページの画面データや、他のプログラム又はコマンドなどのデータをネットワークインターフェース部69を介して対応するロボット1に送出する。

【0048】このようにしてコンテンツサーバ61においては、アクセスしてきたロボット1に対してWebページの画面データや、この他の必要なデータを送受信することができるようになされている。

【0049】なおコンテンツサーバ61内のハードディスク装置68内にはそれぞれ複数のデータベース(図示せず)が格納されており、各種処理を実行するときに対応するデータベースから必要な情報を読み出し得るようになされている。

【0050】このうちのデータベースには、なぞなぞ等の言葉遊びに必要な大量のコンテンツデータが格納されている。かかるコンテンツデータには、言葉遊びに使用する実際の内容を表すデータに加えて、当該言葉遊びに付随して得られる種々の内容を表すオプションデータが付加されている。

【0051】例えば言葉遊びとして「なぞなぞ」が指定された場合、コンテンツデータは「なぞなぞ」の問題及びその解答並びにその理由を表し、当該コンテンツデータに付加されたオプションデータは、当該問題の難易度やその問題が出題された回数から得られる人気の指標等を表す。

【0052】そしてロボット1は、ユーザとの対話において、マイクロホン51を介して集音したユーザの発話内容を後述する音声認識処理を実行することにより認識し、当該認識結果をユーザに関連する種々のデータと共にネットワーク62を介してコンテンツサーバ61に送信する。

【0053】続いてコンテンツサーバ61は、ロボット1から得られた認識結果等に基づいて、データベースに格納されている大量のコンテンツデータの中から最適なコンテンツデータを抽出し、当該コンテンツデータを元のロボット1の送信する。

【0054】かくしてロボット1は、コンテンツサーバ61から取得したコンテンツデータに基づく音声をスピーカ54を介して放音することにより、あたかも人間同士で対話しているかのごとく、自然な感じでユーザと「なぞなぞ」の言葉遊びをすることができるようになされている。

【0055】(3) 名前学習機能に関するメイン制御部

40の処理

次にこのロボット1に搭載された名前学習機能について説明する。

【0056】このロボット1には、人との対話を通してその人の名前を取得し、当該名前を、マイクロホン51の出力に基づいて検出したその人の声の音響的特徴のデータと関連付けて記憶すると共に、これら記憶した各データに基づいて、名前を取得していない新規な人の登場を認識し、その新規な人の名前や声の音響的特徴を上記と同様にして取得し記憶するようにして、人の名前をその人と対応付けて取得(以下、これを名前の学習と呼ぶ)学習していく名前学習機能が搭載されている。なお以下においては、その人の声の音響的特徴と対応付けて名前を記憶し終えた人を『既知の人』と呼び、記憶し終えていない人を『新規な人』と呼ぶものとする。

【0057】そしてこの名前学習機能は、メイン制御部40における各種処理により実現されている。

【0058】ここで、かかる名前学習機能に関するメイン制御部40の処理内容を機能的に分類すると、図8に示すように、人が発声した言葉を認識する音声認識部80と、人の声の音響的特徴を検出すると共に当該検出した音響的特徴に基づいてその人を識別して認識する話者認識部81と、人との対話制御を含む新規な人の名前学習のための各種制御や、既知の人の名前及び声の音響的特徴の記憶管理を司る対話制御部82と、対話制御部82の制御のもとに各種対話用の音声信号S3を生成してスピーカ54(図5)に送出する音声合成部83とに分けることができる。

【0059】この場合、音声認識部80においては、マイクロホン51(図5)からの音声信号S1Bに基づき所定の音声認識処理を実行することにより当該音声信号S1Bに含まれる言葉を単語単位で認識する機能を有するものであり、認識したこれら単語を文字列データD1として対話制御部82に送出するようになされている。

【0060】また話者認識部81は、マイクロホン51から与えられる音声信号S1Bに含まれる人の声の音響的特徴を、例えば“Segregation of Speakers for Recognition and Speaker Identification (CH2977-7/91/00 00~0873 \$1.00 1991 IEEE)”に記載された方法等を利用した所定の信号処理により検出する機能を有している。

【0061】そして話者認識部81は、通常時には、この検出した音響的特徴のデータをそのとき記憶している全ての既知の人の音響的特徴のデータと順次比較し、そのとき検出した音響的特徴がいずれか既知の人の音響的特徴と一致した場合には当該既知の人の音響的特徴と対応付けられた当該音響的特徴に固有の識別子(以下、これをSIDと呼ぶ)を対話制御部82に通知する一方、検出した音響的特徴がいずれの既知の人の音響的特徴とも一致しなかった場合には、認識不能を意味するSID

(=1)を対話制御部82に通知するようになされている。

【0062】また話者認識部81は、対話制御部82が新規な人であると判断したときに当該対話制御部82から与えられる新規学習の開始命令及び学習終了命令に基づいて、その間その人の声の音響的特徴を検出し、当該検出した音響的特徴のデータを新たな固有のSIDと対応付けて記憶すると共に、このSIDを対話制御部82に通知するようになされている。

【0063】なお話者認識部81は、対話制御部82からの追加学習の開始命令及び終了命令に応じて、その人の声の音響的特徴のデータを追加的に収集する追加学習を行い得るようになされている。

【0064】音声合成部83は、対話制御部82から与えられる文字列データD2を音声信号S3に変換する機能を有し、かくして得られた音声信号S3をスピーカ54(図5)に送出するようになされている。これによりこの音声信号S3に基づく音声をスピーカ54から出力させることができるようになされている。

【0065】対話制御部82においては、図9に示すように、既知の人の名前と、話者認識部81が記憶しているその人の声の音響的特徴のデータに対応付けられたSIDとを関連付けて記憶するメモリ84(図8)を有している。

【0066】そして対話制御部82は、所定のタイミングで所定の文字列データD2を音声合成部83に与えることにより、話し相手の人に対して名前を質問し又は名前を確認するための音声等をスピーカ54から出力させる一方、このときのその人の応答等に基づく音声認識部80及び話者認識部81の各認識結果と、メモリ84に格納された上述の既知の人の名前、SIDの関連付けの情報とに基づいてその人が新規な人であるか否かを判断するようになされている。

【0067】そして対話制御部82は、その人が新規な人であると判断したときには、話者認識部81に対して新規学習の開始命令及び終了命令を与えることにより、これら話者認識部81にその新規な人の声の音響的特徴のデータを収集及び記憶させると共に、この結果としてこれら話者認識部81から与えられるその新規な人の声の音響的特徴のデータに対応付けられたSIDを、かかる対話により得られたその人の名前と関連付けてメモリ84に格納するようになされている。

【0068】また対話制御部82は、その人が既知の人であると判断したときには、必要に応じて話者認識部81に追加学習の開始命令を与えることにより話者認識部81に追加学習を行わせる一方、これと共に音声合成部83に所定の文字列データD2を所定のタイミングで順次送出することにより、話者認識部81が追加学習をするのに必要な相当量のデータを収集できるまでその人との対話を長引かせるような対話制御を行うようになされ

ている。

【0069】(4)名前学習機能に関する対話制御部82の具体的処理

次に、名前学習機能に関する対話制御部82の具体的な処理内容について説明する。

【0070】対話制御部82は、外部メモリ58(図5)に格納された制御プログラムに基づいて、図10及び図11に示す名前学習処理手順RT1に従って新規な人の名前を順次学習するための各種処理を実行する。

【0071】すなわち対話制御部82は、マイクロホン51からの音声信号S1Bに基づき話者認識部81が人の声の音声的特徴を認識することにより当該話者認識部81からSIDが与えられると名前学習処理手順RT1をステップSP0において開始し、続くステップSP1において、メモリ84に格納された既知の人の名前と、これに対応するSIDとを関連付けた情報(以下、これを関連付け情報と呼ぶ)に基づいてそのSIDから対応する名前を検索できるか否か(すなわちSIDが認識不能を意味する「-1」でないか否か)を判断する。

【0072】ここでこのステップSP1において肯定結果を得ることは、その人が、話者認識部81がその人の声の音声的特徴のデータを記憶しており、当該データと対応付けられたSIDがその人の名前と関連付けてメモリ84に格納されている既知の人であることを意味する。ただしこの場合においても、話者認識部81が新規の人を既知の人と誤認識したことも考えられる。

【0073】そこで対話制御部82は、ステップSP1において肯定結果を得た場合には、ステップSP2に進んで所定の文字列データD2を音声合成部83に送出することにより、例えば図12に示すように、「〇〇さんですよ。」といったその人の名前がSIDから検索された名前(上述の〇〇に当てはまる名前)と一致するか否かを確かめるための質問の音声をスピーカ54から出力させる。

【0074】次いで対話制御部82は、ステップSP3に進んで、かかる質問に対するその人の「はい、そうです。」や「いいえ、違います。」といった応答の音声認識結果が音声認識部80から与えられるのを待ち受ける。そして対話制御部82は、やがて音声認識部80からかかる音声認識結果が与えられ、また話者認識部81からそのときの話者認識結果であるSIDが与えられると、ステップSP4に進んで、音声認識部80からの音声認識結果に基づき、その人の応答が肯定的なものであるか否かを判断する。

【0075】ここでこのステップSP4において肯定結果を得ることは、ステップSP1において話者認識部81から与えられたSIDに基づき検索された名前がその人の名前と一致しており、従ってその人は対話制御部82が検索した名前を有する本人であるとほぼ断定できる状態にあることを意味する。

【0076】かくしてこのとき対話制御部82は、その人は当該対話制御部82が検索した名前を有する本人であると断定し、ステップSP5に進んで話者認識部61に対して追加学習の開始命令を与える。

【0077】そして対話制御部82は、この後ステップSP6に進んで例えば図12のように「今日はいい天気ですね。」などといった、その人との対話を長引かせるための雑談をさせるための文字列データD2を音声合成部83に順次送出し、この後追加学習に十分な所定時間が経過すると、ステップSP7に進んで話者認識部81に対して追加学習の終了命令を与えた後、ステップSP20に進んでその人に対する名前学習処理を終了する。

【0078】一方、ステップSP1において否定結果を得ることは、話者認識部81により声認識された人が新規の人であるか、又は話者認識部81が既知の人を新規の人と誤認識したことを意味する。またステップSP4において否定結果を得ることは、最初に話者認識部81から与えられたSIDから検索された名前がその人の名前と一致していないことを意味する。そして、これらいずれの場合においても、対話制御部82がその人を正しく把握していない状態にあるといえる。

【0079】そこで対話制御部82は、ステップSP1において否定結果を得たときや、ステップSP4において否定結果を得たときには、ステップSP8に進んで音声合成部83に文字列データD2を与えることにより、例えば図13に示すように、「あれ、名前を教えてください。」といった、その人の名前を聞き出すための質問の音声をスピーカ54から出力させる。

【0080】そして対話制御部82は、この後ステップSP9に進んで、かかる質問に対するその人の「〇〇です。」といった応答の音声認識結果（すなわち名前）と、当該応答時における話者認識部81の話者認識結果（すなわちSID）とがそれぞれ音声認識部80及び話者認識部81から与えられるのを待ち受ける。

【0081】そして対話制御部82は、やがて音声認識部80から音声認識結果が与えられ、話者認識部81からSIDが与えられると、ステップSP10に進んで、これら音声認識結果及びSIDに基づいて、その人が新規な人であるか否かを判断する。

【0082】ここでこの実施の形態の場合、かかる判断は、音声認識部80の音声認識により得られた名前と、話者認識部81からのSIDとでなる2つの認識結果の多数決により行われ、いずれか一方でも否定的な認識結果が得られれば保留することとする。

【0083】例えば、話者認識部81からのSIDが認識不能を意味する「-1」で、かつステップSP9において音声認識部80からの音声認識結果に基づき得られたその人の名前がメモリ84においてどのSIDとも関連付けられていない場合には、その人が新規な人であると判断する。既知のどの顔又はどの声とも似つかない人

が全く新しい名前をもっているという状況であるので、そのような判断ができる。

【0084】また対話制御部82は、話者認識部81からのSIDがメモリ84において異なる名前と関連付けられており、かつステップSP9において音声認識部80からの音声認識結果に基づき得られたその人の名前がメモリ84に格納されてない場合にも、その人が新規な人であると判断する。これは、各種認識処理において、新規カテゴリを既知カテゴリのどれかと誤認識するのは起こり易いことであり、また音声認識された名前が登録されていないことを考えれば、かなり高い確信度をもって新規の人と判断できるからである。

【0085】これに対して対話制御部82は、話者認識部81からのSIDがメモリ84において同じ名前と関連付けられており、かつステップSP9において音声認識部80からの音声認識結果に基づき得られたその人の名前がそのSIDが関連付けられた名前である場合には、その人が既知の人であると判断する。

【0086】また対話制御部82は、話者認識部81からのSIDがメモリ84において異なる名前と関連付けられており、かつステップSP9において音声認識部80からの音声認識結果に基づき得られたその人の名前がかかるSIDが関連付けられた名前である場合には、その人が既知の人であるか又は新規の人であるかを判断しない。このケースでは、音声認識部80及び話者認識部81のいずれか又は両方の認識が間違っていることも考えられるが、この段階ではそれを判定することができない。従ってこの場合には、かかる判断を保留する。

【0087】そして対話制御部82は、このような判断処理により、ステップSP10において、かかる人が新規の人であると判断した場合には、ステップSP11に進んで新規学習の開始命令を話者認識部81に与え、この後ステップSP12に進んで例えば図13のように「私はロボットです。よろしくお願いします。」又は「〇〇さん、今日はいい天気ですね。」などのその人との対話を長引かせる雑談をするための文字列データD2を音声合成部83に送出する。

【0088】また対話制御部82は、この後ステップSP13に進んで話者認識部81における音響的特徴のデータの収集が十分量に達したか否かを判断し、否定結果を得るとステップSP12に戻って、この後ステップSP13において肯定結果を得るまでステップSP12-SP13-SP12のループを繰り返す。

【0089】そして対話制御部82は、やがて話者認識部81における音響的特徴のデータの収集が十分量に達することによりステップSP13において肯定結果を得ると、ステップSP14に進んで、これら話者認識部81に新規学習の終了命令を与える。この結果、話者認識部81において、その音響的特徴のデータが新たなSIDと対応付けられて記憶される。

【0090】また対話制御部82は、この後ステップSP15に進んで、話者認識部81からかかるSIDが与えられるのを待ち受け、やがてこれが与えられると、例えば図14に示すように、これらをステップSP9において音声認識部80からの音声認識結果に基づき得られたその人の名前と関連付けてメモリ84に登録する。そして対話制御部82は、この後ステップSP20に進んでその人に対する名前学習処理を終了する。

【0091】これに対して対話制御部82は、ステップSP10において、かかる人が既知の人であると判断した場合には、ステップSP16に進んで、話者認識部81がその既知の人を正しく認識できていた場合（すなわち話者認識部81が、関連付け情報としてメモリ84に格納されたその既知の人に対応するSIDと同じSIDを認識結果として出力していた場合）には、その話者認識部81に対して追加学習の開始命令を与える。

【0092】具体的には、対話制御部82は、ステップSP9において得られた話者認識部61からのSIDと、最初に話者認識部81から与えられたSIDとがメモリ84において同じ名前と関連付けられており、かつステップSP9において音声認識部80からの音声認識結果に基づき得られた名前がそのSIDが関連付けられた名前であることによりステップSP10においてその人が既知の人であると判断したときには、話者認識部81に対して追加学習の開始命令を与える。

【0093】そして対話制御部82は、この後ステップSP17に進んで、例えば図15に示すように、「ああ〇〇さんですね。思い出しましたよ。今日はいい天気ですね。」「前はえーと、いつ会いましたっけ。」などのその人との対話を長引かせるための雑談をさせるための文字列データD2を音声合成部83に順次送出し、この後追加学習に十分な所定時間が経過すると、ステップSP18に進んで話者認識部81に対して追加学習の終了命令を与えた後、ステップSP20に進んでその人に対する名前学習処理を終了する。

【0094】また話者認識部81は、ステップSP9において得られた話者認識部81からのSIDと、最初に話者認識部81から与えられたSIDとがメモリ65において異なる名前と関連付けられており、かつステップSP9において音声認識部80からの音声認識結果に基づき得られた名前がかかるSIDが関連付けられた名前であることによりステップSP10においてその人が既知の人であるとも新規の人であるとも判定できないと判断した場合、ステップSP19に進んで、例えば図16に示すように、「ああそうですか。元気ですか。」などの雑談をさせるための文字列データD2を音声合成部83に順次送出する。

【0095】そしてこの場合には、対話制御部82は、新規学習又は追加学習の開始命令及びその終了命令を話者認識部81に与えず（すなわち新規学習及び追加学習

のいずれも話者認識部81に行わせず）、所定時間が経過すると、ステップSP20に進んでその人に対する名前学習処理を終了する。

【0096】このようにして対話制御部82は、音声認識部80及び話者認識部81の各認識結果に基づいて、人との対話制御や話者認識部81の動作制御を行うことにより、新規な人の名前を順次学習することができるようになされている。

【0097】このようにこのロボット1では、新規な人との対話を通してその人の名前を取得し、当該名前を、マイクロホン51の出力に基づいて検出したその人の声の音響的特徴のデータと関連付けて記憶すると共に、これら記憶した各種データに基づいて、名前を取得していないさらに新規な人の登場を認識し、その新規な人の名前や声の音響的特徴及び顔の形態的特徴を上述と同様にして取得し記憶するようにして、人の名前を学習することができる。

【0098】従って、このロボット1は、音声コマンドの入力やタッチセンサの押圧操作等のユーザからの明示的な指示による名前登録を必要とすることなく、人間が普段行うように、通常の人との対話を通して新規な人物や物体等の名前を自然に学習することができる。

【0099】（5）音声認識部80の具体的構成次に、図17において、上述のような名前学習機能を具現化するための音声認識部80の具体的構成について説明する。

【0100】この音声認識部80においては、マイクロホン51からの音声信号S1BをAD（Analog Digital）変換部90に入力する。AD変換部90は、供給されるアナログ信号である音声信号S1Bをサンプリング、量子化し、ディジタル信号である音声データにA/D変換する。この音声データは、特徴抽出部91に供給される。

【0101】特徴抽出部91は、そこに入力される音声データについて、適当なフレームごとに、例えば、MFCC（Mel Frequency Cepstrum Coefficient）分析を行い、その分析の結果得られるMFCCを、特徴ベクトル（特徴パラメータ）として、マッチング部92と未登録語区間処理部96に出力する。なお、特徴抽出部91では、その後、例えば線形予測係数、ケプストラム係数、線スペクトル対、所定の周波数ごとのパワー（フィルタバンクの出力）等を、特徴ベクトルとして抽出することが可能である。

【0102】マッチング部92は、特徴抽出部91からの特徴ベクトルを用いて、音響モデル記憶部93、辞書記憶部94及び文法記憶部95を必要に応じて参照しながら、マイクロホン51に入力された音声（入力音声）を、例えば、連続分布HMM（Hidden Markov Model）法に基づいて音声認識する。

【0103】すなわち音響モデル記憶部93は、音声認

識する音声の言語における個々の音素や、音節、音韻などのサブワードについて音響的な特徴を表す音響モデル（例えば、HMMの他、DP（Dynamic Programing）マッチングに用いられる標準パターン等を含む）を記憶している。なお、ここでは連続分布HMM法に基づいて音声認識を行うことをしているので、音響モデルとしてはHMM（Hidden Markov Model）が用いられる。

【0104】辞書記憶部94は、認識対象の各单位ごとにクラスタリングされた、その単語の発音に関する情報（音響情報）と、その単語の見出しとが対応付けられた単語辞書を認識している。

【0105】ここで、図18は、辞書記憶部94に記憶された単語辞書を示している。

【0106】図18に示すように、単語辞書においては、単語の見出しとその音韻系列とが対応付けられており、音韻系列は、対応する単語ごとにクラスタリングされている。図18の単語辞書では、1つのエントリ（図16の1行）が、1つのクラスに相当する。

【0107】なお、図18において、見出しはローマ字と日本語（仮名漢字）で表してあり、音韻系列はローマ字で表してある。ただし、音韻系列における「N」は、撥音「ん」を表す。また、図18では、1つのエントリに1つの音韻系列を記述してあるが、1つのエントリには複数の音韻系列を記述することも可能である。

【0108】図17に戻り、文法記憶部95は、辞書記憶部94の単語辞書に登録されている各単語がどのように連鎖する（つながる）かを記述した文法規則を記憶している。

【0109】ここで、図19は、文法記憶部95に記憶された文法規則を示している。なお、図19の文法規則は、EBNF（Extended Backus Naur Form）で記述されている。

【0110】図19においては、行頭から最初に現れる「;」までが1つの文法規則を表している。また先頭に「\$」が付されたアルファベット（列）は変数を表し、「\$」が付されていないアルファベット（列）は単語の見出し（図18に示したローマ字による見出し）を表す。さらに「[]」で囲まれた部分は省略可能であることを表し、「|」は、その前後に配置された見出しの単語（あるいは変数）のうちのいずれか一方を選択することを表す。

【0111】従って、図19において、例えば、第1行（上から1行目）の文法規則「\$col = [Kono | sono] ir o wa ;」は、変数\$colが、「このいろ（色）」または「そのいろ（色）」という単語列であることを表す。

【0112】なお、図19に示した文法規則においては、変数\$silと\$garbageが定義されていないが、変数\$silは、無音の音響モデル（無音モデル）を表し、変数\$garbageは、基本的には、音韻どうしの間での自由な遷移を許可したガーベジモデルを表す。

【0113】再び図17に戻り、マッチング部92は、辞書記憶部94の単語辞書を参照することにより、音響モデル記憶部93に記憶されている音響モデルを接続することで、単語の音響モデル（単語モデル）を構成する。さらにマッチング部92は、幾つかの単語モデルを文法記憶部95に記憶された文法規則を参照することにより接続し、そのようにして接続された単語モデルを用いて、特徴ベクトルに基づき、連続分布HMM法によって、マイクロホン51に入力された音声認識する。すなわちマッチング部92は、特徴抽出部91が出力する時系列の特徴ベクトルが観測されるスコア（尤度）が最も高い単語モデルの系列を検出し、その単語モデルの系列に対応する単語列の見出しを、音声の認識結果として出力する。

【0114】より具体的には、マッチング部92は、接続された単語モデルに対応する単語により接続し、そのようにして接続された単語モデルを用いて、特徴ベクトルに基づき、連続分布HMM法によって、マイクロホン51に入力された音声認識する。すなわちマッチング部92は、特徴抽出部91が出力する時系列の特徴ベクトルが観測されるスコア（尤度）が最も高い単語モデルの系列を検出し、その単語モデルの系列に対応する単語列の見出しを音声認識結果として出力する。

【0115】より具体的には、マッチング部92は、接続された単語モデルに対応する単語列について、各特徴ベクトルの出現確率（出力確率）を累積し、その累積値をスコアとして、そのスコアを最も高くする単語列の見出しを音声認識結果として出力する。

【0116】以上のようにして出力されるマイクロホン51に入力された音声認識結果は、文字列データD1として対話制御部82に出力される。

【0117】ここで図19の実施の形態では、第9行（上から9行目）にガーベジモデルを表す変数\$garbageを用いた文法規則（以下、適宜、未登録語用規則という）「\$pat1 = \$color | \$garbage \$color2 ;」があるが、マッチング部92は、この見登録語用規則が適用された場合には、変数\$garbageに対応する音声区間を未登録語の音声区間として検出する。さらに、マッチング部92は、未登録語用規則が適用された場合における変数\$garbageが表すガーベジモデルにおける音韻の遷移としての音韻系列を未登録語の音韻系列として検出する。そしてマッチング部92は、未登録語用規則が適用された音声認識結果が得られた場合に検出される未登録語の音声区間と音韻系列を未登録語区間処理部96に供給する。

【0118】なお上述の未登録語用規則「\$pat1 = \$color | \$garbage \$color2 ;」によれば、変数#color1で表される単語辞書に登録されている単語（列）の音韻系列と、変数\$color2で表される単語辞書に登録されている単語（列）の音韻系列との間にある1つの未登録語が検

出されるが、この実施の形態においては、発話に複数の未登録語が含まれている場合や、未登録語が単語辞書に登録されている単語（列）間に挟まれていない場合であっても適用可能である。

【0119】未登録語区間処理部96は、特徴抽出部91から供給される特徴ベクトルの系列（特徴ベクトル系列）を一時記憶する。さらに、未登録語区間処理部96は、マッチング部92から未登録語の音声区間と音韻系列を受信すると、その音声区間における音声の特徴ベクトル系列を、一時記憶している特徴ベクトル系列から検出する。そして未登録語区間処理部96は、マッチング部92からの音韻系列（未登録語）にユニークなID（identification）を付し、未登録語の音韻系列と、その音声区間における特徴ベクトル系列とともに、特徴ベクトルバッファ97に供給する。

【0120】特徴ベクトルバッファ97は、例えば、図20に示すように、未登録語区間処理部96から供給される未登録語のID、音韻系列及び特徴ベクトル系列を対応付けて一時記憶する。

【0121】ここで図20においては、未登録語に対して1からのシーケンシャルな数値がIDとして付されている。従って、例えばいま、特徴ベクトルバッファ97において、N個の未登録語のID、音韻系列及び特徴ベクトル系列が記憶されている場合において、マッチング部92が未登録語の音声区間と音韻系列を検出すると、未登録語区間処理部96では、その未登録語に対してN+1がIDとして付され、特徴ベクトルバッファ97では、図20に点線で示すように、その未登録語のID、音韻系列及び特徴ベクトル系列が記憶される。

【0122】再び図17に戻り、クラスタリング部98は、特徴ベクトルバッファ97に新たに記憶された未登録語（以下、適宜、新未登録語という）について、特徴ベクトルバッファ97に既に記憶されている他の未登録語（以下、適宜、既記憶未登録語という）それぞれに対するスコアを計算する。

【0123】すなわちクラスタリング部98は、新未登録語を入力音声とし、かつ既記憶未登録語を単語辞書に登録されている単語とみなして、マッチング部92における場合と同様にして、新未登録語について、各既記憶未登録語に対するスコアを計算する。具体的には、クラスタリング部98は、特徴ベクトルバッファ97を参照することで新未登録語の特徴ベクトル系列を認識するとともに、既記憶未登録語の音韻系列にしたがって音響モデルを接続し、その接続された音響モデルから新未登録語の特徴ベクトル系列が観測される尤度としてのスコアを計算する。

【0124】なお、音響モデルは、音響モデル記憶部93に記憶されているものが用いられる。

【0125】クラスタリング部98は、同様にして、各既記憶未登録語について、新未登録語に対するスコアも

計算し、そのスコアによってスコアシート記憶部99に記憶されたスコアシートを更新する。

【0126】さらにクラスタリング部98は、更新したスコアシートを参照することにより、既に求められている未登録語（既記憶未登録語）をクラスタリングしたクラスタの中から、新未登録語を新たなメンバとして加えるクラスタを検出する。さらにクラスタリング部98は、新未登録語を検出したクラスタの新たなメンバとし、そのクラスタをそのクラスタのメンバに基づいて分割し、その分割結果に基づいて、スコアシート記憶部99に記憶されているスコアシートを更新する。

【0127】スコアシート記憶部99は、新未登録語についての既記憶未登録語に対するスコアや、既記憶未登録語についての新未登録語に対するスコア等が登録されたスコアシートを記憶する。

【0128】ここで、図21は、スコアシートを示している。

【0129】スコアシートは、未登録語の「ID」、「音韻系列」、「クラスタナンバ」、「代表メンバID」及び「スコア」が記述されたエントリで構成される。

【0130】未登録語の「ID」と「音韻系列」としては、特徴ベクトルバッファ97に記憶されたものと同一のものがクラスタリング部98によって登録される。「クラスタナンバ」は、そのエントリの未登録語がメンバとなっているクラスタを特定するための数字で、クラスタリング部98によって付され、スコアシートに登録される。「代表ナンバID」は、そのエントリの未登録語がメンバとなっているクラスタを代表する代表メンバとしての未登録のIDであり、この代表メンバIDによって、未登録語がメンバとなっているクラスタの代表メンバを認識することができる。なお、クラスタの代表メンバは、クラスタリング部98によって求められ、その代表メンバのIDがスコアシートの代表メンバIDに登録される。「スコア」は、そのエントリの未登録語についての他の未登録語それぞれに対するスコアであり、上述したように、クラスタリング部98によって計算される。

【0131】例えば、いま、特徴ベクトルバッファ97において、N個の未登録語のID、音韻系列及び特徴ベクトル系列が記憶されているとすると、スコアシートには、そのN個の未登録語のID、音韻系列、クラスタナンバ、代表ナンバID及びスコアが登録されている。

【0132】そして特徴ベクトルバッファ97に、新未登録語のID、音韻系列、および特徴ベクトル系列が新たに記憶されると、クラスタリング部98では、スコアシートが図21において点線で示すように更新される。

【0133】すなわちスコアシートには、新未登録語のID、音韻系列、クラスタナンバ、代表メンバID、新未登録語についての既記憶未登録語それぞれに対するスコア（図19におけるスコア $s(N+1, 1)$ 、 $s(2, N+1)$ 、 $\dots s(N+1, N)$ ）が追加される。さら

にスコアシートには、既記憶未登録語それぞれについての未登録語に対するスコア（図21における $s(N+1, 1)$ 、 $s(2, N+1)$ 、 $\dots s(N+1, N)$ ）が追加される。さらに後述するように、スコアシートにおける未登録語のクラスタナンバと代表メンバIDが必要に応じて変更される。

【0134】なお、図21の実施の形態においては、IDが i の未登録語（の発話）についての、IDが j の未登録語（の音韻系列）に対するスコアを、 $s(i, j)$ として表してある。

【0135】またスコアシート（図21）には、IDが i の未登録語（の発話）についての、IDが i の未登録語（の音韻系列）に対するスコア $s(i, j)$ も登録される。ただし、このスコア $s(i, j)$ は、マッチング部92において、未登録語の音韻系列が検出されるときに計算されるため、クラスタリング部98で計算する必

$$K = \max_k \left\{ \sum s(k^3, k) \right\} \dots\dots (1)$$

【0139】で示される値 $k (\in k)$ をIDとするメンバが代表メンバとされることになる。

【0140】ただし、(1)式において、 $\max_k \{ \}$ は、 $\{ \}$ 内の値を最大にする k を意味する。また k^3 は、 k と同様に、クラスタに属するメンバのIDを意味する。さらに、 \sum は、 k^3 をクラスタに属するメンバすべてのIDに亘って変化させての総和を意味する。

【0141】なお上述のように代表メンバを決定する場合、クラスタのメンバが1または2つの未登録語であるときには、代表メンバを決めるにあたってスコアを計算する必要はない。すなわちクラスタのメンバが1つの未登録語である場合には、その1つの未登録語が代表メンバとなり、クラスタのメンバが2つの未登録語である場合には、その2つの未登録語のうちのいずれを代表メンバとしても良い。

【0142】また代表メンバの決定方法は、上述したものに限定されるものではなく、その他、例えばクラスタのメンバとなっている未登録語のうち、他の未登録語それぞれとの特徴ベクトル空間における距離の総和を最小にするもの等をそのクラスタの代表メンバとすることも可能である。

【0143】以上のように構成される音声認識部80では、マイクロホン51に入力された音声認識する音声認識処理と、未登録語に関する未登録語処理が図22に示す音声認識処理手順RT2に従って行われる。

【0144】實際上、音声認識部80では、人が発話を行うことにより得られた音声信号S1Bがマイクロホン51からAD変換部90を介して音声データとされて特徴抽出部91に与えられるとこの音声認識処理手順RT2がステップSP30において開始される。

【0145】そして続くステップSP31において、特徴抽出部91が、その音声データを所定のフレーム単位

要はない。

【0136】再び図17に戻り、メンテナンス部100は、スコアシートに記憶部99における更新後のスコアシートに基づいて、辞書記憶部94に記憶された単語辞書を更新する。

【0137】ここで、クラスタの代表メンバは、次のように決定される。すなわち、例えば、クラスタのメンバとなっている未登録語のうち、他の未登録語それぞれについてのスコアの総和（その他、例えば、総和を他の未登録語の数で除算した平均値でも良い）を最大にするものがそのクラスタの代表メンバとされる。従って、この場合、クラスタに属するメンバのメンバIDを k で表すこととすると、次式

【0138】

【数1】

で音響分析することにより特徴ベクトルを抽出し、その特徴ベクトルの系列をマッチング部92及び未登録語区間処理部96に供給する。

【0146】マッチング部96は、続くステップS32において、特徴抽出部91からの特注オベクトル系列について、上述したようにスコア計算を行い、この後ステップS33において、スコア計算の結果得られるスコアに基づいて、音声認識結果となる単語列の見出しを求めて出力する。

【0147】さらにマッチング部92は、続くステップS34において、ユーザの音声に未登録語が含まれていたかどうかを判定する。

【0148】ここで、このステップS34において、ユーザの音声に未登録語が含まれていないと判定された場合、すなわち上述の未登録語用規則「\$pat1=\$color| \$garbage \$color2:」が適用されずに音声認識結果が得られた場合、ステップS35に進んで処理が終了する。

【0149】これに対してステップS34において、ユーザの音声に未登録語が含まれていると判定された場合、すなわち未登録語用規則「\$pat1=\$color| \$garbage \$color2:」が適用されて音声認識結果が得られた場合、マッチング部92は、続くステップS35において、未登録語用規則の変数\$garbageに対応する音声区間を未登録語の音声区間として検出するとともに、その変数\$garbageが表すガーベジモデルにおける音韻の遷移としての音韻系列を未登録語の音韻系列として検出し、その未登録語の音声区間と音韻系列を未登録語区間処理部96に供給して、処理を終了する（ステップSP36）。

【0150】一方、未登録語機関処理部96は、特徴抽出部91から供給される特徴ベクトル系列を一時記憶し

ており、マッチング部92から未登録語の音声区間と音韻系列が供給されると、その音声区間における音声の特徴ベクトル系列を検出する。さらに未登録語区間処理部96は、マッチング部92からの未登録語（の音韻系列）にIDを付し、未登録語の音韻系列と、その音声区間における特徴ベクトル系列とともに、特徴ベクトルバッファ97に供給する。

【0151】以上のようにして、特徴ベクトルバッファ97に新たな未登録語（新未登録語）のID、音韻系列及び特徴ベクトル系列が記憶されると、この後、未登録語の処理が図23に示す未登録語処理手順RT3に従って行われる。

【0152】すなわち音声認識部80においては、上述のように特徴ベクトルバッファ97に新たな未登録語（新未登録語）のID、音韻系列及び特徴ベクトル系列が記憶されるとこの未登録語処理手順RT3がステップSP40において開始され、まず最初にステップS41において、クラスタリング部98が、特徴ベクトルバッファ97から新未登録語のIDと音韻系列を読み出す。

【0153】次いでステップS42において、クラスタリング部98が、スコアシート記憶部99のスコアシートを参照することにより、既に求められている（生成されている）クラスタが存在するかどうかを判定する。

【0154】そしてこのステップS42において、すでに求められているクラスタが存在しないと判定された場合、すなわち新未登録語が初めての未登録語であり、スコアシートに既記憶未登録語のエントリが存在しない場合には、ステップS43に進み、クラスタリング部98が、その新未登録語を代表メンバとするクラスタを新たに生成し、その新たなクラスタに関する情報と、親身登録語に関する情報とをスコアシート記憶部99のスコアシートに登録することにより、スコアシートを更新する。

【0155】すなわちクラスタリング部98は、特徴ベクトルバッファ97から読み出した新未登録語のIDおよび音韻系列をスコアシート（図21）に登録する。さらにクラスタリング部98は、ユニークなクラスタナンバを生成し、新未登録語のクラスタナンバとしてスコアシートに登録する。またクラスタリング部98は、新未登録語のIDをその新未登録語の代表ナンバIDとして、スコアシートに登録する。従ってこの場合は、新未登録語は、新たなクラスタの代表メンバとなる。

【0156】なお、いまの場合、新未登録語とのスコアを計算する既記憶未登録語が存在しないため、スコアの計算は行われない。

【0157】かかるステップS43の処理後は、ステップS52に進み、メンテナンス部100は、ステップS43で更新されたスコアシートに基づいて、辞書記憶部94の単語辞書を更新し、処理を終了する（ステップSP54）。

【0158】すなわち、いまの場合、新たなクラスタが生成されているので、メンテナンス部100は、スコアシートにおけるクラスタナンバを参照し、その新たに生成されたクラスタを認識する。そしてメンテナンス部100は、そのクラスタに対応するエントリを辞書記憶部94の単語辞書に追加し、そのエントリの音韻系列として、新たなクラスタの代表メンバの音韻系列、つまりいまの場合は、新未登録語の音韻系列を登録する。

【0159】一方、ステップS42において、すでに求められているクラスタが存在すると判定された場合、すなわち新未登録語が初めての未登録語ではなく、従ってスコアシート（図21）に、既記憶未登録語のエントリ（行）が存在する場合、ステップS44に進み、クラスタリング部98は、新未登録語について、各既記憶未登録語それぞれに対するスコアを計算すると共に、各既記憶未登録語それぞれについて、新未登録語に対するスコアを計算する。

【0160】すなわち、例えば、いま、IDが1乃至N個の既記憶未登録語が存在し、新未登録語のIDをN+1とすると、クラスタリング部98では、図21において点線で示した部分の新未登録語についてのN個の既記憶未登録語それぞれに対するスコア $s(N+1, 1)$ 、 $s(N+1, 2) \dots, s(N, N+1)$ と、N個の既記憶未登録語それぞれについての新未登録語に対するスコア $s(1, N+1)$ 、 $s(2, N+1) \dots, s(N, N+1)$ が計算される。なおクラスタリング部98において、これらのスコアを計算するにあたっては、新未登録語とN個の既記憶未登録語それぞれの特徴ベクトル系列が必要となるが、これらの特徴ベクトル系列は、特徴ベクトルバッファ97を参照することで認識される。

【0161】そしてクラスタリング部98は、計算したスコアを新未登録語のID及び音韻系列とともにスコアシート（図21）に追加し、ステップS45に進む。

【0162】ステップS45では、クラスタリング部98はスコアシート（図21）を参照することにより、新未登録語についてのスコア $s(N+1, i)$ （ $i=1, 2, \dots, N$ ）を最も高く（大きく）する代表メンバを有するクラスタを検出する。即ち、クラスタリング部98は、スコアシートの代表メンバIDを参照することにより、代表メンバとなっている既記憶未登録語を認識し、さらにスコアシートのスコアを参照することで、新未登録語についてのスコアを最も高くする代表メンバとしての既記憶未登録語を検出する。そしてクラスタリング部98は、その検出した代表メンバとしての既記憶未登録語のクラスタナンバのクラスタを検出する。

【0163】その後、ステップS46に進み、クラスタリング部98は、新未登録語をステップS45で検出したクラスタ（以下、適宜、検出クラスタという）のメンバに加える。すなわちクラスタリング部98は、スコアシートにおける新未登録語のクラスタナンバとして、検

出クラスタの代表メンバのクラスタナンバを書き込む。

【0164】そしてクラスタリング部98は、ステップS47において、検出クラスタを例えば2つのクラスタに分割するクラスタ分割処理を行い、ステップS48に進む。ステップS48では、クラスタリング部98は、ステップS47のクラスタ分割処理によって、検出クラスタを2つのクラスタに分割することができたかどうか判定し、分割することができた判定した場合、ステップS49に進む。ステップS49では、クラスタリング部98は、検出クラスタの分割により得られる2つのクラスタ（この2つのクラスタを、以下、適宜、第1の子

$$D(k1, k2) = \max_k \left\{ \text{abs}(\log(s(k, k1)) - \log(s(k, k2))) \right\} \dots\dots (2)$$

【0168】で表される値 $D(k1, k2)$ を第1と第2の子クラスタ同士の間の子クラスタ間距離とする。

【0169】ただし、(2)式において、 $\text{abs}()$ は、()内の値の絶対値を表す。また、 $\max_k \{ \}$ は、 k を変えて求められる $\{ \}$ 内の値の最大値を表す。また \log は、自然対数又は常用対数を表す。

【0170】いま、 ID が i のメンバをメンバ# i と表すこととすると、(2)式におけるスコアの逆数 $1/s(k, k1)$ は、メンバ# k と代表メンバ $k1$ との距離に相当し、スコアの逆数 $1/s(k, k2)$ は、メンバ# k と代表メンバ $k2$ との距離に相当する。従って、

(2)式によれば、第1と第2の子クラスタのメンバのうち、第1の子クラスタの代表メンバ# $k1$ との距離と、第2の子クラスタの代表メンバ# $k2$ との差の最大値が、第1と第2の子クラスタ同士の間の子クラスタ間距離とされることになる。

【0171】なおクラスタ間距離は、上述したものに限定されるものではなく、その他、例えば、第1の子クラスタの代表メンバと、第2の子クラスタの代表メンバとのDPマッチングを行うことにより、特徴ベクトル空間における距離の積算値を求め、その距離の積算値を、クラスタ間距離とすることも可能である。

【0172】ステップS49の処理後は、ステップS50に進み、クラスタリング部98は、第1と第2の子クラスタ同士の子クラスタ間距離が、所定の閾値より大である（あるいは、閾値以上である）かどうかを判定する。

【0173】ステップS50において、クラスタ間距離が所定の閾値より大であると判定された場合、すなわち検出クラスタのメンバとしての複数の未登録後が、その音響的特徴からいって、2つのクラスタにクラスタリングすべきものであると考えられる場合、ステップS51に進み、クラスタリング部98は、第1と第2の子クラスタをスコアシート記憶部99のスコアシートに登録する。

【0174】すなわちクラスタリング部98は、第1と第2の子クラスタにユニークなクラスタナンバを割り当

ラスタと第2の子クラスタという）同士の間の子クラスタ間距離を求める。

【0165】ここで、第1及び第2の子クラスタ同士間の子クラスタ間距離とは、例えば次のように定義される。

【0166】すなわち第1の子クラスタと第2の子クラスタの両方の任意のメンバ（未登録語）の ID を、 k で表すとともに、第1と第2の子クラスタの代表メンバ（未登録語）の ID を、それぞれ $k1$ または $k2$ で表すこととすると、次式

【0167】

【数2】

て、検出クラスタのメンバのうち、第1の子クラスタにクラスタリングされたもののクラスタナンバを第1の子クラスタのクラスタナンバにすると共に、第2の子クラスタにクラスタリングされたもののクラスタナンバを第2の子クラスタのクラスタナンバにするように、スコアシートを更新する。

【0175】さらにクラスタリング部98は、第1の子クラスタにクラスタリングされたメンバの代表メンバ ID を第1の子クラスタの代表メンバの ID にすると共に、第2の子クラスタにクラスタリングされたメンバの代表メンバ ID を第2の子クラスタの代表メンバの ID にするように、スコアシートを更新する。

【0176】なお、第1と第2の子クラスタのうちいずれか一方には、検出クラスタのクラスタナンバを割り当てることができる。

【0177】クラスタリング部98が以上のようにして第1と第2の子クラスタをスコアシートに登録すると、ステップS51からS52に進み、メンテナンス部100が、スコアシートに基づいて、辞書記憶部94の単語辞書を更新し、処理を終了する（ステップSP54）。

【0178】すなわち、いまの場合、検出クラスタが第1と第2の子クラスタに分割されたため、メンテナンス部100は、まず単語辞書における検出クラスタに対応するエントリを削除する。さらにメンテナンス部100は、第1と第2の子クラスタそれぞれに対応する2つのエントリを単語辞書に追加し、第1の子クラスタに対応するエントリの音韻系列として、その第1の子クラスタの代表メンバの音韻系列を登録すると共に、第2の子クラスタに対応するエントリの音韻系列として、その第2の子クラスタの代表メンバの音韻系列を登録する。

【0179】一方、ステップS48において、ステップS47のクラスタ分割処理によって、検出クラスタを2つのクラスタに分割することができなかったと判定された場合、又はステップS50において、第1と第2の子クラスタの子クラスタ間距離が所定の閾値より大でないと判定された場合、従って、検出クラスタのメンバとしての複数の未登録後の音響的特徴が第1と第2の子ク

スタにクラスタリングするほど似ていないものではない場合)、ステップS53に進み、クラスタリング部98は、検出クラスタの新たな代表メンバを求め、スコアシートを更新する。

【0180】すなわちクラスタリング部98は、新未登録後をメンバとして加えた検出クラスタの各メンバについて、スコアシート記憶部99のスコアシートを参照することにより、(1)式の計算に必要なスコア $s(k, k)$ を認識する。さらに、クラスタリング部98は、その認識したスコア $s(k, k)$ を用い、(1)式に基づき、検出クラスタの新たな代表メンバとなるメンバのIDを求める。そしてクラスタリング部98は、スコアシート(図21)における検出クラスタの各メンバの代表メンバIDを、検出クラスタの新たな代表メンバのIDに書き換える。

【0181】その後、ステップS52に進み、メンテナンス部100が、スコアシートに基づいて辞書記憶部94の単語辞書を更新し、処理を終了する(ステップSP54)。

【0182】すなわち、いまの場合、メンテナンス部100は、スコアシートを参照することにより、検出クラスタの新たな代表メンバを認識し、さらにその代表メンバの音韻系列を認識する。そしてメンテナンス部100は、単語辞書における検出クラスタに対応するエントリの音韻系列を、検出クラスタの新たな代表メンバの音韻系列に変更する。

【0183】ここで、図23のステップSP47のクラスタ分割処理は、図24に示すクラスタ分割処理手順RT4に従って行われる。

【0184】すなわち音声認識部80では、図24のステップSP46からステップSP47に進むとこのクラスタ分割処理手順RT4をステップSP60において開始し、まず最初にステップS61において、クラスタリング部98が、新未登録後がメンバとして加えられた検出クラスタから、まだ選択していない任意の2つのメンバの組み合わせを選択し、それぞれを仮の代表メンバとする。ここで、この2つの仮の代表メンバを、以下、適宜、第1の仮代表メンバと第2の仮代表メンバという。

【0185】そして、続くステップS62において、クラスタリング部98は、第1の仮代表メンバ及び第2の仮代表メンバをそれぞれ代表メンバとすることができるように、検出クラスタのメンバを2つのクラスタに分割することができるかどうかを判定する。

【0186】ここで、第1又は第2の仮代表メンバを代表メンバとすることができるかどうかは(1)式の計算を行う必要があるが、この計算に用いられるスコア $s(k', k)$ は、スコアシートを参照することで認識される。

【0187】ステップS62において、第1の仮代表メンバ及び第2の仮代表メンバをそれぞれ代表メンバとす

ることができるように、検出クラスタのメンバを2つのクラスタに分割することができないと判定された場合、ステップS62をスキップして、ステップS64に進む。

【0188】また、ステップS62において、第1の仮代表メンバと、第2の仮代表メンバをそれぞれ代表メンバとすることができるように、検出クラスタのメンバを2つのクラスタに分割することができると判定された場合、ステップS63に進み、クラスタリング部98は、第1の仮代表メンバと、第2の仮代表メンバがそれぞれ代表メンバとなるように、検出クラスタのメンバを2つのクラスタに分割し、その分割後の2つのクラスタの組を、検出クラスタの分割結果となる第1及び第2の子クラスタの候補(以下、適宜、候補クラスタの組という)として、ステップS64に進む。

【0189】ステップS64では、クラスタリング部98は、検出クラスタのメンバの中で、まだ第1と第2の仮代表メンバの組として選択していない2つのメンバの組があるかどうかを判定し、あると判定した場合、ステップS61に戻り、まだ第1と第2の仮代表メンバの組として選択していない検出クラスタの2つのメンバの組が選択され、以下、同様の処理が繰り返される。

【0190】またステップS64において、第1と第2の仮代表メンバの組として選択していない検出クラスタの2つのメンバの組がないと判定された場合、ステップS65に進み、クラスタリング部98は、候補クラスタの組が存在するかどうかを判定する。

【0191】ステップS65において、候補クラスタの組が存在しないと判定された場合、ステップS66をスキップして、リターンする。この場合は、図23のステップS48において、検出クラスタを分割することができなかったと判定される。

【0192】一方、ステップS65において、候補クラスタの組が存在すると判定された場合、ステップS66に進み、クラスタリング部98は、候補クラスタの組が複数存在するときには、各候補クラスタの組の2つのクラスタ同士の間のクラスタ間距離を求める。そして、クラスタリング部98は、クラスタ間距離が最小の候補クラスタの組を求め、その候補クラスタの組を検出クラスタの分割結果をして、すなわち第1と第2の子クラスタとして、リターンする。なお、候補クラスタの組が1つだけの場合は、その候補クラスタの組がそのまま第1と第2の子クラスタとされる。

【0193】この場合は、図23のステップS48において、検出クラスタを分割することができたと判定される。

【0194】以上のように、クラスタリング部98において、既に求められている未登録語をクラスタリングしたクラスタの中から、新未登録語を新たなメンバとして加えるクラスタ(検出クラスタ)を検出し、新未登録語

をその検出クラスタの新たなメンバとして、検出クラスタをその検出クラスタのメンバに基づいて分割するようにしたので、未登録語をその音響的特徴が近似しているもの同士に容易にクラスタリングすることができる。

【0195】さらにメンテナンス部100において、そのようなクラスタリング結果に基づいて単語辞書を更新するようにしたので、単語辞書の大規模化を避けながら、未登録語の単語辞書への登録を容易に行うことができる。

【0196】また、例えば、仮に、マッチング部92において、未登録語の音声区間の検出を誤ったとしても、そのような未登録語は、検出クラスタの分割によって、音声区間が正しく検出された未登録語とは別のクラスタにクラスタリングされる。そして、このようなクラスタに対応するエントリが単語辞書に登録されることになるが、このエントリの音韻系列は正しく検出されなかった音声区間に対応するものとなるから、その後の音声認識において大きなスコアを与えることはない。従って、仮に、未登録語の音声区間の検出を誤ったとしても、その誤りはその後の音声認識にはほとんど影響しない。

【0197】ここで、図25は、未登録語の発話を行って得られたクラスタリング結果を示している。なお、図25においては、各エントリ（各行）が1つのクラスタを表している。また、図25の左欄は、各クラスタの代表メンバ（未登録語）の音韻系列を表しており、図25の右欄は、各クラスタのメンバとなっている未登録語の発話内容と数を表している。

【0198】すなわち図25において、例えば第1行のエントリは、未登録語「風呂」の1つの発話だけがメンバとなっているクラスタを表しており、その代表メンバの音韻系列は、「doroa:」（ドロア）になっている。また、例えば第2行のエントリは、未登録語「風呂」の3つの発話がメンバとなっているクラスタを表しており、その代表メンバの音韻系列は、「kuroj」（クロ）になっている。

【0199】さらに、例えば第7行のエントリは、未登録語「本」の4つの発話がメンバとなっているクラスタを表しており、その代表メンバの音韻系列は、「NhoNde:su」（ンホンテース）になっている。また、例えば第8行のエントリは、未登録語「オレンジ」の1つの発話と、未登録語「本」の19の発話がメンバとなっているクラスタを表しており、その代表メンバの音韻系列は、「ohoN」（オホン）になっている。他のエントリも同様のことを表している。

【0200】図25によれば、同一の未登録語の発話について、良好にクラスタリングされていることが分かる。

【0201】なお、図25の第8行のエントリにおいては、未登録語「オレンジ」の1つの発話と、未登録語「本」の19の発話が、同一のクラスタにクラスタリングされ

ている。このクラスタはそのメンバとなっている発話から、未登録語「本」のクラスタとなるべきであると考えられるが、未登録語「オレンジ」の発話も、そのクラスタのメンバとなっている。しかしながらこのクラスタも、その後に未登録語「本」の発話がさらに入力されていくと、クラスタ分割され、未登録語「本」の発話だけをメンバとするクラスタと、未登録語「オレンジ」の発話だけをメンバとするクラスタにクラスタリングされると考えられる。

【0202】（5）対話制御システムを用いたユーザとロボットとの対話

（5-1）言葉遊びにおけるコンテンツデータの取得及び提供

実際に図6に示す対話制御システム63では、ユーザがロボット1との間で言葉遊びによる対話を行う場合、ユーザからの要求に応じてロボット1が言葉遊びの具体的な内容（例えば「なぞなぞ」）を表すコンテンツデータをコンテンツサーバ61内のデータベースから取得して、当該コンテンツデータに基づく問題等をユーザに対して発話することができるようになされている。

【0203】この対話制御システムにおいて、ロボット1は、ユーザから例えば「なぞなぞをしよう」という発話をスピーカ54を介して集音すると、図26に示すコンテンツデータ取得処理手順RT5をステップSP70から開始し、続くステップSP71において、ユーザの発話内容を音声認識処理した後、ユーザごとに対応して作成しておいたプロフィールデータをメイン制御部40内のメモリ40Aから読み出してロードする。

【0204】かかるプロフィールデータは、メイン制御部40内のメモリ40Aに格納されており、図27に示すように、ユーザごとに既に行った言葉遊びの種類が記述され、さらに当該種類ごとにそれぞれ問題の難易度（レベル）、既に遊んだID及び当該遊んだ回数が記述されている。

【0205】具体的には、まずユーザ名が「○田△子」のユーザでは、言葉遊びのうち「なぞなぞ」について、レベルが「2」、既に遊んだIDが「1、3、…」及び遊んだ回数が「10」であり、「山手線ゲーム」について、レベルが「4」、既に遊んだIDが「1、2、…」及び遊んだ回数が「5」である。またユーザ名が「□山×男」のユーザでは、言葉遊びのうち「なぞなぞ」について、レベルが「5」、既に遊んだIDが「3、4、…」及び遊んだ回数が「30」であり、「山手線ゲーム」について、レベルが「2」、既に遊んだIDが「2、5、…」及び遊んだ回数が「2」である。

【0206】そしてこのプロフィールデータは、コンテンツサーバ61に送出する一方、当該コンテンツサーバ61からフィードバックされることにより適宜更新されるようになされている。具体的には、言葉遊びのうち「なぞなぞ」について、正解すれば難易度（レベル）を上げると共に、人気があれば面白くない問題であったと

判断してそのタイプの問題を避けるようにプロフィールデータを更新する。

【0207】そしてロボット1は、ステップSP72において、言葉遊びのうち「なぞなぞ」を要求するデータをネットワーク62を介してコンテンツサーバ61に送信した後、ステップSP73に進む。

【0208】コンテンツサーバ61は、ロボット1から要求データを受信すると、コンテンツデータ提供処理手順RT6をステップSP80から開始し、続くステップSP81において、該当するロボット1との間で通信可能な接続状態を確立する。

【0209】ここでコンテンツサーバ61内のデータベースには、言葉遊びの種類（例えば「なぞなぞ」や「山手線ゲーム」等）ごとにコンテンツデータが生成され、当該コンテンツデータは、その種類に合わせて設定された複数の出題内容がID番号を付して記述されている。

【0210】例えば図28に示すように、言葉遊びのうち「なぞなぞ」について4個の出題内容が順次ID番号が割り当てられて記述されている（以下、これらを第1～第4の出題内容ID1～ID4という）。これら第1～第4の出題内容ID1～ID4は、それぞれ問題と、当該問題に対する答えと、当該答えに対する理由とが順次記述されたものである。

【0211】まず第1の出題内容ID1では、問題が「4歳と5歳の子供しか住んでいない外国の都市は？」、答えが「シカゴ」、及び理由が「4歳と5歳でシカゴだよ」として記述されている。また第2の出題内容ID2では、問題が「少ししか人が乗っていないのに一杯な車はなんだ？」、答えが「救急車」、及び理由が「キュウキュウで一杯だよ」として記述されている。さらに第3の出題内容ID3では、問題が「家の中で暖房が効かない場所はどこだ？」、答えが「玄関」、及び理由が「厳しい寒さで厳寒だよ」として記述されている。さらに第4の出題内容ID4では、問題が「落ち込んでいても2回食べると元気になるのは？」、答えが「海苔」、及び理由が「2回でのりのりだよ」として記述されている。

【0212】そしてコンテンツデータには、言葉遊びの種類に応じて設定されるオプションデータが付加されており、第1～第4の出題内容ID1～ID4に対応してそれぞれ問題の難易度及び出題回数に応じた人気度が数値化されて記述されている。このオプションデータはロボット1からのアクセス回数やユーザの解答結果等に基づいて内容が逐次更新されるようになされている。

【0213】続いてコンテンツサーバ61は、ロボット1に対して「なぞなぞ」についてのコンテンツデータに付加されたオプションデータを送信した後、ステップSP83に進む。

【0214】やがてロボットは、ステップSP73において、コンテンツサーバ61から送信されたオプションデータを受信すると、当該オプションデータとユーザに

対応するプロフィールデータとを比較する。そしてロボット1は、コンテンツデータの中から該当するユーザに最も合った出題内容を選択して、当該出題内容を要求する旨のデータをネットワーク62を介してコンテンツサーバ61に送信する。

【0215】具体的には上述した図27に示すように、例えばユーザ名が「○田△子」のユーザが言葉遊びのうち「なぞなぞ」をする場合、このユーザについてのプロフィールデータをコンテンツサーバ61に送信して、当該プロフィールデータに基づく「なぞなぞ」のレベル「2」に相当する出題内容を表すコンテンツデータを要求する。

【0216】コンテンツサーバ61は、ステップSP83において、ロボット1から送信されたデータに基づいて、データベースから対応するコンテンツデータを読み出した後、ネットワーク62を介してロボット1に送信し、ステップSP84に進む。

【0217】具体的にはロボット1から得られたプロフィールデータが「なぞなぞ」のレベルが「2」を表す場合、そのレベルに合った問題、すなわち図28に示すオプションデータのうち難易度「2」に相当する出題内容を表すコンテンツデータを選択してロボット1に送信する。この場合、コンテンツデータのうち第1及び第4の出題内容ID1、ID4が該当するが、ユーザ名「○田△子」における既に遊んだIDが「1」を含むため、第1の出題内容ID1ではなく、未だ遊んだことのない第4の出題内容ID4をコンテンツサーバ61はロボット1に送信する。

【0218】そしてステップSP74において、ロボット1は、コンテンツサーバ61から取得したコンテンツデータをロードした後、ステップSP75に進んで、コンテンツサーバ61に対して通信接続の切断要求を表す旨のデータをネットワーク62を介して送信し、ステップSP76に進んで当該コンテンツデータ取得処理手順RT5を終了する。

【0219】一方、コンテンツサーバ61は、ステップSP84において、ロボット1から送信されたデータに基づいて、当該ロボット1との間で確立されている通信接続を切断した後、ステップSP85に進んで当該コンテンツデータ提供処理手順RT6を終了する。

【0220】このようにしてコンテンツデータ取得処理手順RT5においては、ロボット1は、ユーザと言葉遊びをする際、当該言葉遊びのうちユーザによって特定の種類（なぞなぞ等）が指定されたとき、当該種類を構成する複数の出題内容の中からユーザに最適な出題内容をコンテンツサーバ61から取得することができる。

【0221】またコンテンツデータ提供処理手順RT6においては、コンテンツサーバ61は、ロボット1からの要求に応じて、データベースに格納されている複数のコンテンツデータのうちユーザに最適な出題内容を含むコンテンツデータを選択してロボット1に提供すること

ができる。

【0222】(5-2) ロボットとユーザとの言葉遊びによる対話シーケンス

ここでロボット1のメイン制御部40内のメモリ40Aには、ロボット1とユーザとが言葉遊びによる対話を行う場合に、当該言葉遊びの種類ごとに、ロボット1とユーザとの対話のやり取りを表す対話モデルが予め決められており、当該対話モデルに基づいて、言葉遊びの種類が同一であれば(例えば「なぞなぞ」に関する限り)、コンテンツデータを入れ替えるだけで、新たに異なる出題内容等をユーザに提供することができるようになされている。

【0223】実際にロボット1はユーザから言葉遊びを行う旨の発話を受け取ると、図29に示すように、ロボット1のメモリ制御部40がこの言葉遊びの種類に対応する対話モデルに基づいて、ユーザとの対話のときに次のロボット1による発話内容を順次決定していくようになされている。

【0224】かかる対話モデルでは、ロボット1がとり得る発話をそれぞれノードNDB1~NDB7として、遷移可能なノード間を発話を表す有向アークで結び、かつ1つのノード間で完結する発話を自己発話アークとして表現する有向グラフを用いる。

【0225】このためメモリ40Aには、このような有向グラフの元となる、当該ロボット1が発話できる全ての発話をデータベース化したファイルが格納されており、このファイルに基づいて有向グラフを生成する。

【0226】ロボット1のメイン制御部40は、ユーザから言葉遊びを行う旨の発話を受け取ると、対応する有向グラフを用いて、有向アークの向きに従いながら現在のノードから指定された発話に対応付けられた有向アーク若しくは自己動作アークに至る経路を探索し、当該探索した経路上の各有向アークにそれぞれ対応付けられた発話を順次行わせるような指令を次々と出力するようになされている。

【0227】実際にユーザとロボット1との間で言葉遊びの種類のうち「なぞなぞ」による対話を行う場合を説明する。まずロボット1が例えば「4歳と5歳の子供しか住んでいない外国の都市はどこでしょう?」という出題内容を表すコンテンツデータをコンテンツサーバ61から取得して(ノードND1)、当該出題内容をユーザに向けて発話する(ノードND2)。

【0228】そしてロボット1はユーザからの応答を待ち(ノードND3)、ユーザの発話が正解である「シカゴ」であれば、「あたり〜!」と発話して(ノードND4)、その理由である「4と5でシカゴだよ」と発話する(ノードND7)。

【0229】またユーザの発話が不正解であれば、「ちがうよ。答え聞く?」と発話した後(ノードND5)、ユーザから「はい」という返事が得られれば「答えはね

え、シカゴ!」と答えを発話した後(ノードND6)、さらにその理由である「4と5でシカゴだよ」と発話する(ノードND7)一方、「きかない」という返事が得られれば、再度ユーザからの応答を待つ(ノードND3)。

【0230】さらにユーザの発話が「こうさん」であれば、「答えはねえ、シカゴ!」と答えを発話した後(ノードND6)、さらにその理由である「4と5でシカゴだよ」と発話する(ノードND7)。また一定時間が経過しても、ユーザから何も発話されないときには、ロボット1は「ねえねえ、まだ?」と発話して(ノードND3)、ユーザからの応答を促すようにする。

【0231】このようにロボット1はユーザの発話に関連する応答として、単に正解を発話するのみならず、正解の理由をも発話することにより、ユーザにとってロボット1と「なぞなぞ」をするときの面白さを増大させることができる。

【0232】さらにこのように正解の理由をもロボット1が発話することにより、ロボット1がユーザの発話内容を誤認識した場合でもそのことをユーザは知ることができる。

【0233】これはゲームなので、ユーザがあえてロボット1の音声認識の誤りを訂正する必要は特にはないが、ロボット1がユーザの発話内容を誤認識した場合でも、それを間接的にユーザに伝えることで、言葉遊びのゲームをスムーズに進行させることができる。

【0234】(5-3) オプションデータの更新
図6に示す対話制御システム63では、上述したコンテンツデータ取得処理手順RT5及びコンテンツデータ提供処理手順RT6(図26)において述べたように、ロボット1がコンテンツサーバ61からコンテンツデータを取得すると、どのデータを取得したのかの情報がそのコンテンツデータに付加されたオプションデータに反映される。

【0235】例えば、ロボット1が言葉遊びのうち何の種類さらには何の出題内容を何回取得したかの指標となる人気のデータの値が変更される。

【0236】またロボット1がユーザに言葉遊びを出題したときに、その出題内容に対してユーザが正解したか否かのデータも、ネットワーク62を介してコンテンツサーバ61にフィードバックされ、当該問題の難易度に反映されるようにその値が更新される。

【0237】このようにロボット1からコンテンツサーバ61内のデータベースへのフィードバックは、ユーザが意識することなくロボット1によって自動的に行われるものもあるが、例えばロボット1との対話によってコンテンツサーバ61へのフィードバックをユーザから直接取得するようにしても良い。

【0238】ここでコンテンツサーバ61において、ロボット1からフィードバックされたコンテンツデータに

基づいて、当該コンテンツデータに付加されたオプションデータを更新する場合について説明する。

【0239】ロボット1がコンテンツサーバ61からコンテンツデータを取得すると、どのデータを取得したのかの情報がそのコンテンツデータに付加されたオプションデータに反映される。

【0240】実際に図6に示す対話制御システム63では、ユーザがロボット1との間で言葉遊びによる対話を行った後、ロボット1が人気指標を更新すると自発的又はユーザからの発話に応じて決定すると、図30に示す人気指標集計処理手順RT7をステップSP90から開始し、続くステップSP91において、コンテンツサーバ61に対してアクセス要求を表すデータを送信する。

【0241】コンテンツサーバ61は、ロボット1から要求データを受信すると、オプションデータ更新処理手順RT8をステップSP100から開始し、続くステップSP101において、該当するロボット1との間で通信可能な接続状態を確立する。

【0242】そしてロボット1は、ステップSP92に進んで、「今の問題面白かった?」といった質問をユーザに対して発話した後、ステップSP93に進む。

【0243】このステップSP93において、ロボット1は、ユーザからの応答を待った後、当該応答を受け取ったときステップSP94に進む。このステップSP94において、ロボット1は、ユーザからの応答の内容が「つまなかった」又は「おもしろかった」のいずれかを判断し、「つまなかった」と判断した場合にはステップSP95に進んで、人気のレベル値をデクリメント（減少）させるように要求する旨の要求データをネットワーク62を介してコンテンツサーバ61に送信した後、ステップSP97に進む。

【0244】これに対してステップSP94において、ロボット1は、ユーザからの応答の内容が「おもしろかった」と判断した場合にはステップSP96に進んで、人気のレベル値をインクリメント（増加）させるように要求する旨の要求データをネットワーク62を介してコンテンツサーバ61に送信した後、ステップSP97に進む。

【0245】コンテンツサーバ61は、ステップSP102において、ロボット1から送信された要求データに基づいて、データベースから対応するコンテンツデータに付加されたオプションデータを読み出した後、当該オプションデータの記述内容のうち「人気度」の値を減少又は増加させる。

【0246】そしてコンテンツサーバ61は、ステップSP103において、オプションデータの更新が終了した旨の応答データをネットワーク62を介してロボット1に送信した後、ステップSP104に進む。

【0247】ロボット1は、コンテンツサーバ61から送信された応答データに基づいて、オプションデータが

更新された旨を確認した後、当該コンテンツサーバ61に対して通信接続の切断要求を表す旨の要求データをネットワーク62を介してコンテンツサーバ61に送信し、そのままステップSP98に進んで当該人気指標集計処理手順RT7を終了する。

【0248】コンテンツサーバ61は、ステップSP104において、ロボット1から送信された要求データに基づいて、当該ロボット1との間で確立されている通信接続を切断した後、ステップSP105に進んで当該オプションデータ更新処理手順RT8を終了する。

【0249】このようにして人気指標集計処理手順RT7においては、ロボット1は、ユーザに出題したコンテンツデータに基づく出題内容について、その面白さの是非を当該ユーザに問うことにより、その問題の人気の有無を確認することができる。

【0250】またオプションデータ更新処理手順RT8においては、ロボット1から得られたコンテンツデータに基づく出題内容についての人気の有無に基づいて、当該コンテンツデータに付加されたオプションデータの記述内容を更新することにより、そのユーザにとって当該出題内容の面白さや好み等を次回の際に反映させることができる。

【0251】（5-4）コンテンツデータの登録

ここでコンテンツサーバ61内のデータベースに格納されている言葉遊びの種類ごとに登録されているコンテンツデータは、当該コンテンツデータに基づく出題内容及びその答え並びにその答えの理由（以下、単に出題内容等と呼ぶ）を、各ユーザが発話することによりロボット1を介して間接的にコンテンツサーバ61に登録させる場合と、各ユーザがロボット1を介することなく、自己の個人端末等を用いて直接的にコンテンツサーバ61に登録させる場合の2通りがある。以下にそれぞれの場合について説明する。

【0252】（5-4-1）ロボット1を介して間接的に出題内容等を追加登録させる場合

図6に示す対話制御システム63では、ユーザの発話により出題内容等を受け取ったロボット1は、当該出題内容等をネットワーク62を介してコンテンツサーバ61に送信することにより、当該コンテンツデータ内のデータベースに追加登録させるようになされている。

【0253】この対話制御システム63において、ロボット1は、ユーザから新しい出題内容等を表す発話をスピーカ54を介して集音すると、図31に示すコンテンツ収集処理手順RT9をステップSP110から開始し、続くステップSP111において、コンテンツサーバ61に対してアクセス要求を表す要求データを送信する。

【0254】そしてコンテンツサーバ61は、ロボット1から要求データを受信すると、コンテンツデータ追加登録処理手順RT10をステップSP120から開始

し、続くステップSP121において、該当するロボット1との間で通信可能な接続状態を確立する。

【0255】そしてロボット1は、ステップSP112に進んで、ユーザから取得した出題内容等を表す取得データをネットワーク62を介してコンテンツサーバ61に送信した後、ステップSP113に進む。

【0256】コンテンツサーバ61は、ステップSP122において、ロボット1から送信された取得データに基づいて、当該取得データをコンテンツデータとしてID番号を割り当てた後、ステップSP123に進む。

【0257】このステップSP123では、コンテンツサーバ61は、データベースにおいて該当するユーザに対応しかつ言葉遊びの種類に対応する記憶位置に、当該ID番号を割り当てた出題内容等を登録する。この結果、データベースには、該当するユーザにおける言葉遊びの該当する種類において、第N（Nは自然数）の出題内容IDNが追加して記述されることとなる。

【0258】そしてコンテンツサーバ61は、ステップSP124に進んで、コンテンツデータの追加登録が終了した旨の応答データをネットワーク62を介してロボット1に送信した後、ステップSP125に進む。

【0259】ロボット1は、コンテンツサーバ61から送信された応答データに基づいて、コンテンツデータが追加登録された旨を確認した後、当該コンテンツサーバ61に対して通信接続の切断要求を表す旨の要求データをネットワーク62を介してコンテンツサーバ61に送信し、そのままステップSP114に進んで当該コンテンツ収集処理手順RT9を終了する。

【0260】コンテンツサーバ61は、ステップSP125において、ロボット1から送信された要求データに基づいて、当該ロボット1との間で確立されている通信接続を切断した後、ステップSP126に進んで当該コンテンツデータ追加登録処理手順RT10を終了する。

【0261】このようにしてコンテンツ収集処理手順RT9においては、ロボット1は、ユーザから発話した新しい出題内容等を、コンテンツサーバ61内のデータベースにそのユーザに応じたコンテンツデータとして追加登録させることができる。

【0262】またコンテンツデータ追加登録処理手順RT10においては、ロボット1から得られた出題内容等に基づいて、当該出題内容等をコンテンツデータとしてそのユーザに関する記述内容に追加して登録することにより、当該ユーザのみならず他のユーザにとってもコンテンツの種類が増大した分より一層面白さを増すことができる。

【0263】このことは新たな出題内容等を発話したユーザにとっても、コンテンツサーバ61にアクセスしてデータベースに格納されているオプションデータを読み出すことにより、自分が提案した出題内容等がどの程度他のユーザに使用されているかなどを知ることができ、

出題内容等の登録そのものを楽しみを持たせることができる。

【0264】ここで上述した対話モデルを用いて、実際にロボット1がユーザの発話により出題内容等を受け取ると、図31に示すように、ロボット1のメモリ制御部40がこの言葉遊びの種類に対応する対話モデルに基づいて、ユーザとの対話のときに次のロボット1による発話内容を順次決定していくようになされている。

【0265】まずロボット1が「面白い問題教えて」とユーザに向けて発話する。そしてロボット1はユーザからの応答を待ち（ノードND10）、ユーザの発話が「いいよ」であれば、「問題を言ってよ」と発話した後（ノードND11）、さらにユーザからの応答を待つ。

【0266】一方、ユーザの発話が「いやだ」であれば、「う～ん、残念」と発話した後（ノードND12）、かかる対話シーケンスを終了する。

【0267】やがてロボット1は、ユーザから問題として例えば「落ち込んでいても2回食べると元気になる食べ物は何？」という発話を受け取ると、その音声認識結果（問題の言葉）を繰り返し発話する（ノードND13）。

【0268】この発話を聞いたユーザが「そうだよ」と発話した場合には、ロボット1は「答えは？」とその問題の答えを要求する発話を行う一方（ノードND14）、ユーザが「ちがうよ」と発話した場合には、ロボット1は「もう一回問題を言ってよ」と再度問題を要求する発話を行う（ノードND11）。

【0269】そしてユーザから答えである「海苔」という発話を受け取ると、その音声認識結果（答えの言葉）を繰り返し発話する（ノードND15）。この発話を聞いたユーザが「そうだよ」と発話した場合には、ロボット1は「理由は？」とその答えの理由を要求する発話を行う一方（ノードND16）、ユーザが「ちがうよ」と発話した場合には、ロボット1は「もう一回答えを言ってよ」と再度答えを要求する発話を行う（ノードND14）。

【0270】そしてユーザから理由である「2回でノリノリだよ」という発話を受け取ると、その音声認識結果（理由の言葉）を繰り返し発話する（ノードND17）。この発話を聞いたユーザが「そうだよ」と発話した場合には、ロボット1は「じゃ、登録するね」と発話する一方（ノードND18）、ユーザが「ちがうよ」と発話した場合には、ロボット1は「もう一回理由を言ってよ」と再度理由を要求する発話を行う（ノードND16）。

【0271】この後ロボット1はユーザから取得した問題及びその答え並びにその答えの理由をネットワーク62を介してコンテンツサーバ61内のデータベースにコンテンツデータとして追加登録する。

【0272】このようにロボット1は、ユーザから新た

に取得した出題内容等をコンテンツデータとしてそのユーザに関する記述内容に追加して登録することにより、ユーザに対してより一層多くのコンテンツを提供することができる。

【0273】（5-4-2）ロボットを介さずに直接的に出題内容等を修正させる場合

また図6に示す対話制御システム63では、上述のコンテンツ収集処理手順RT9及びコンテンツデータ追加登録処理手順RT10のように、ユーザがロボット1を介して新たな出題内容等をコンテンツサーバ61内のデータベースに追加登録させた後に、ユーザが作成した出題内容等のうち例えば問題の答えに対する理由が、ユーザの発話に関連する応答（すなわち暗に問題の解答の確認）にならない場合や、当該出題内容等の問題が難しくすぎて誰も答えられない場合がある。

【0274】これらの場合には、ユーザが自己のパーソナルコンピュータ等の端末装置を用いてネットワーク62を介してコンテンツサーバ61にアクセスし、データベース内の対応するコンテンツデータの記述内容を修正することができるようになされている。

【0275】具体的には、ユーザが登録した出題内容等について、例えばその問題が「落ち込んでいても2回食べると元気になるのは？」であり、その答え「海苔」に対する理由が「2回食べると元気になるからだよ」となっている場合には、答えである「海苔」を連想させることができない。

【0276】このためコンテンツサーバ61は、ユーザから「理由がよくわからない」などのフィードバックを受けると、ユーザが自己の端末装置を用いてデータベースにアクセスして当該コンテンツデータに基づく出題内容等のうちの理由を「2回でのりのりだよ」と書き換えることで、当該コンテンツデータを修正することができる。

【0277】なお、コンテンツデータの修正は、データベースにアクセスできるユーザのみならず、データベースの管理者が修正しても良い。さらに部分的にコンテンツデータを更新するのみならず、コンテンツデータを全て作成し直すようにしても良い。

【0278】（6）本実施の形態の動作及び効果
以上の構成において、この対話制御システム63では、ロボット1とユーザとの間で言葉遊びによる対話をする際、ユーザから言葉遊びの種類（なぞなぞ等）が指定されたとき、ロボットは、当該ユーザについてのプロフィールデータを読み出して、ネットワーク62を介してコンテンツサーバ61に送信する。

【0279】コンテンツサーバは、ロボット1から受信したプロフィールデータに基づいて、データベースに格納されている複数のコンテンツデータの中からユーザに最適な出題内容等を含むコンテンツデータを選択した後、当該コンテンツデータをロボット1に提供することができる。

【0280】その際、ロボット1とユーザとの言葉遊びの際に、ロボットが発話した出題内容についてユーザが答えた後、ロボットがその答えの理由を一言述べるようにしたことにより、対話自体が知的に見えてより面白くさせることができるのみならず、ロボットがどう認識したのかをユーザに提示することとなり、ユーザが自己の発話と同じである場合にはユーザに安心感を与えることができる一方、ユーザが自己の発話と異なる場合にもその旨をユーザに認識させることができる。

【0281】このようにロボット1がユーザの発話内容をいちいち確認しないため、ユーザとの会話の流れやリズムを止めることがなく、あたかも人間同士が会話しているかのごとく自然な日常会話を実現することができる。

【0282】また対話制御システム63では、ロボット1は、ユーザに出題したコンテンツデータに基づく出題内容について、その面白さの是非を当該ユーザに問いかけ、その結果をコンテンツサーバにフィードバックさせるようにしたことにより、当該コンテンツサーバではその出題内容の人気の有無等について統計的な評価をとることができる。

【0283】さらにコンテンツサーバは、その出題内容についての統計的な評価に基づいて、コンテンツデータに付加されたオプションデータの記述内容を更新することにより、当該ユーザのみならず他のユーザにとってもその出題内容の面白さや好み等を次回の際に反映させることができる。

【0284】さらに対話制御システム63では、ロボット1がユーザから新たに取得した出題内容等をコンテンツサーバに送信して、当該コンテンツサーバにおいてデータベースに追加して登録するようにしたことにより、ユーザに対してより一層多くのコンテンツを提供することができ、その分ユーザに飽きさせることなくロボットとの対話を広く普及させることができる。

【0285】以上の構成によれば、この対話制御システム63において、ロボット1とユーザとの間で言葉遊びによる対話をする際、ユーザから言葉遊びの種類（なぞなぞ等）が指定されたとき、ロボットは当該ユーザについてのプロフィールデータをコンテンツサーバ61に送信し、当該コンテンツサーバ61がデータベースからユーザに最適な出題内容等を含むコンテンツデータを選択してロボット1に提供するようにしたことにより、ロボットの会話に面白みを持たせることができ、かくしてエンターテインメント性を格段的に向上させることができる。

【0286】（7）他の実施の形態

なお上述のように本実施の形態においては、本発明を図1～図3のように構成された2足歩行型のロボット1に適用するようにした場合について述べたが、本発明はこれに限らず、例えば4脚歩行型のロボットなど、この他

種々の形態のペットロボットに広く適用することができる。

【0287】また上述の実施の形態においては、ロボット1において、人間と対話するための機能を有し、当該対話を通じて対象とするユーザの発話を認識する対話手段として、胴体部ユニット2内のメイン制御部40（対話制御部82）を適用するようにした場合について述べたが、本発明はこれに限らず、この他種々の構成からなる対話手段に広く適用するようにしても良い。

【0288】さらに上述の実施の形態においては、ロボット1において、ユーザの発話内容のうち、言葉遊びに関するプロファイルデータ（履歴データ）を生成する生成手段と、当該生成されたプロファイルデータ（履歴データ）を、言葉遊びを通じて得られるユーザの発言内容に応じて更新する更新手段とを、メイン制御部40から構成すると共に、当該プロファイルデータ（履歴データ）をメイン制御部40内のメモリ40Aに格納しておくようにした場合について述べたが、本発明はこれに限らず、生成手段及び更新手段は一体又は別体にかかわらずこの他種々の構成のものに広く適用するようにしても良い。

【0289】また言葉遊びとして、本実施の形態においては、なぞなぞや山手線ゲームを適用したが、これ以外にも、尻取り、しゃれ、語呂あわせ、アナグラム（言葉の綴りの順番を変えて別の意味にする遊び）及び早口言葉など、要するに言葉の発音・リズム・意味などを利用した種々の遊びに広く適用することができる。

【0290】さらに上述の実施の形態においては、ロボット1において、言葉遊びの開始の際には、履歴データをネットワークを介してコンテンツサーバ（情報処理装置）61に送信する通信手段として、胴体部ユニット2内に装着した所定の無線通信規格に対応した無線LANカード（図示せず）を適用するようにした場合について述べたが、本発明はこれに限らず、その他の無線通信回線網のみならず、一般公衆回線やLAN等の有線通信回線網をも適用するようにしても良い。

【0291】さらに上述の実施の形態においては、コンテンツサーバ（情報処理装置）61において、複数の言葉遊びの内容を表すコンテンツデータ（内容データ）を記憶する記憶手段として、コンテンツサーバ61内のハードディスク装置68に格納されたデータベースを適用するようにした場合について述べたが、本発明はこれに限らず、コンテンツデータ（内容データ）を、必要に応じて複数のロボット1が共有できるようにデータベース管理することができれば、種々の構成からなる記憶手段に広く適用するようにしても良い。

【0292】さらに上述の実施の形態においては、コンテンツサーバ（情報処理装置）61において、ロボット1からネットワーク62を介して送信されたプロファイルデータ（履歴データ）を検出する検出手段として、C

PU65を適用するようにした場合について述べたが、本発明はこれに限らず、この他種々の構成の検出手段を適用するようにしても良い。

【0293】さらに上述の実施の形態においては、コンテンツサーバ（情報処理装置）において、検出したプロファイルデータ（履歴データ）に基づいて、データベース（記憶手段）からコンテンツデータ（内容データ）を選択的に読み出してネットワーク62を介して元のロボット1に送信する通信制御手段として、CPU65及びネットワークインターフェイス部69を適用するようにした場合について述べたが、本発明はこれに限らず、この他種々の構成の通信制御手段を適用するようにしても良い。

【0294】さらに上述の実施の形態においては、ロボット1では、ユーザに出力したコンテンツデータ（内容データ）に基づく言葉遊びの内容に関する評価を当該ユーザの発話から認識した後、プロファイルデータ（履歴データ）を評価に応じて更新し、当該更新されたプロファイルデータ（履歴データ）をコンテンツサーバ（情報処理装置）61に送信する。そしてコンテンツサーバ（情報処理装置）61では、言葉遊びのコンテンツデータ（内容データ）に付随するオプションデータ（付随データ）を当該コンテンツデータ（内容データ）に関連付けてデータベース（記憶手段）に記憶しておき、選択されたコンテンツデータ（内容データ）に付随するオプションデータ（付随データ）について、プロファイルデータ（履歴データ）に基づく評価に関連するデータ部分を更新するようにした場合について述べたが、本発明はこれに限らず、要は、オプションデータ（付随データ）を更新することで、コンテンツデータ（内容データ）を当該ユーザのみならず他のユーザにとってもその出題内容の面白さや好み等を次回の際に反映させることができれば、付随データとして他のデータを用いても良く、その更新方法も種々の方法を適用するようにしても良い。

【0295】さらに上述のように本実施の形態においては、ロボットでは、ユーザに出力した新たな言葉遊びの内容を当該ユーザの発話から認識した後、言葉遊びの内容を表す新規内容データをコンテンツサーバ（情報処理装置）61に送信する。そしてコンテンツサーバ（情報処理装置）61では、対応するユーザについての内容データに追加して、新規内容データをデータベース（記憶手段）に記憶するようにした場合について述べたが、本発明はこれに限らず、要は、ユーザに対してより一層多くのコンテンツを提供することによって、その分ユーザに飽きさせることなくロボット1との対話を広く普及させることができれば、新規内容データの追加方法としては他の方法を用いるようにしても良い。

【0296】

【発明の効果】 上述のように本発明によれば、ロボット及び情報処理装置がネットワークを介して接続された対

話制御システムにおいて、ロボットとユーザとの間で言葉遊びによる対話をする際、ユーザの発話内容のうち言葉遊びに関する履歴データを生成して情報処理装置に送信し、当該情報処理装置が記憶手段から当該履歴データに基づいてユーザに最適な内容データを選択的に読み出して元のロボットに提供するようにしたことにより、ユーザとの間でロボットの会話に面白みやリズムを持たせることができ、あたかも人間同士が会話しているかのごとく自然な日常会話に近づけることができ、かくしてエンターテインメント性を格段的に向上させ得る対話制御システムを実現できる。

【0297】また本発明によれば、ロボット及び情報処理装置がネットワークを介して接続された対話制御方法において、ロボットとユーザとの間で言葉遊びによる対話をする際、ユーザの発話内容のうち言葉遊びに関する履歴データを生成して情報処理装置に送信し、当該情報処理装置が履歴データに基づいてユーザに最適な内容データを複数の内容データの中から選択的に読み出して元のロボットに提供するようにしたことにより、ユーザとの間でロボットの会話に面白みやリズムを持たせることができ、あたかも人間同士が会話しているかのごとく自然な日常会話に近づけることができ、かくしてエンターテインメント性を格段的に向上させ得る対話制御方法を実現できる。

【0298】さらに本発明によれば、情報処理装置とネットワークを介して接続されたロボット装置において、人間と対話するための機能を有し、当該対話を通じて対象とするユーザの発話を認識する対話手段と、対話手段によるユーザの発話内容のうち、言葉遊びに関する履歴データを生成する生成手段と、生成手段により生成された履歴データを、言葉遊びを通じて得られるユーザの発言内容に応じて更新する更新手段と、言葉遊びの開始の際には、履歴データをネットワークを介して情報処理装置に送信する通信手段とを設け、情報処理装置において予め記憶された複数の言葉遊びの内容を表す内容データのうち、通信手段から送信された履歴データに基づいて選択された内容データがネットワークを介して送信されたとき、対話手段は、当該内容データに基づく言葉遊びの内容を出力するようにしたことにより、ユーザとの間でロボットの会話に面白みやリズムを持たせることができ、あたかも人間同士が会話しているかのごとく自然な日常会話に近づけることができ、かくしてエンターテインメント性を格段的に向上させ得るロボット装置を実現できる。

【図面の簡単な説明】

【図1】本発明を適用したロボットの外観構成を示す斜視図である。

【図2】本発明を適用したロボットの外観構成を示す斜視図である。

【図3】本発明を適用したロボットの外観構成を示す斜視図である。

【図4】ロボットの内部構成を示すブロック図である。

【図5】ロボットの内部構成を示すブロック図である。

【図6】本実施の形態による対話制御システムの構成を示す略線図である。

【図7】図6に示すコンテンツサーバの構成を示すブロック図である。

【図8】メイン制御部40の処理の説明に供するブロック図である。

【図9】メモリにおけるSIDと名前との関連付けの説明に供する概念図である。

【図10】名前学習処理手順を示すフローチャートである。

【図11】名前学習処理手順を示すフローチャートである。

【図12】名前学習処理時における対話例を示す略線図である。

【図13】名前学習処理時における対話例を示す略線図である。

【図14】SIDと名前との新規登録の説明に供する概念図である。

【図15】名前学習時における対話例を示す略線図である。

【図16】名前学習処理時における対話例を示す略線図である。

【図17】音声認識部の構成を示すブロック図である。

【図18】単語辞書の説明に供する概念図である。

【図19】文法規則の説明に供する概念図である。

【図20】特徴ベクトルバッファの記憶内容の説明に供する概念図である。

【図21】スコアシートの説明に供する概念図である。

【図22】音声認識処理手順を示すフローチャートである。

【図23】未登録語処理手順を示すフローチャートである。

【図24】クラスタ分割処理手順を示すフローチャートである。

【図25】シミュレーション結果を示す概念図である。

【図26】コンテンツデータ取得処理手順及びコンテンツデータ提供処理手順を示すフローチャートである。

【図27】プロフィールデータの説明に供する概念図である。

【図28】コンテンツデータの説明に供する概念図である。

【図29】言葉遊びによる対話シーケンスの説明に供する概念図である。

【図30】人気指標集計処理手順及びオプションデータ更新処理手順を示すフローチャートである。

【図31】コンテンツ収集処理手順及びコンテンツデー

タ追加登録処理手順を示すフローチャートである。

【図32】言葉遊びによる対話シーケンスの説明に供する概念図である。

【符号の説明】

1……ロボット、40……メイン制御部、51……マイクロホン、54……スピーカ、61……コンテンツサーバ、62……ネットワーク、63……対話制御システム、65……CPU、68……ハードディスク装置、6

9……ネットワークインターフェース部、80……音声認識部、81……話者認識部、82……対話制御部、83……音声合成部、84……メモリ、S1B、S3……音声信号、RT5……コンテンツデータ取得処理手順、RT6……コンテンツデータ提供処理手順、RT7……人気指標集計処理手順、RT8……オプションデータ更新処理手順、RT9……コンテンツ収集処理手順、RT10……コンテンツデータ追加登録処理手順。

【図1】

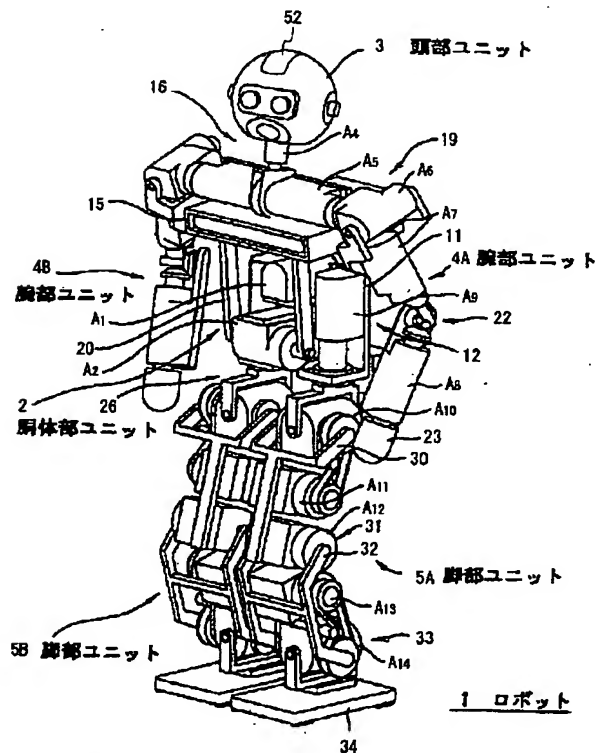


図1 本実施の形態におけるロボットの外觀構成 (1)

【図2】

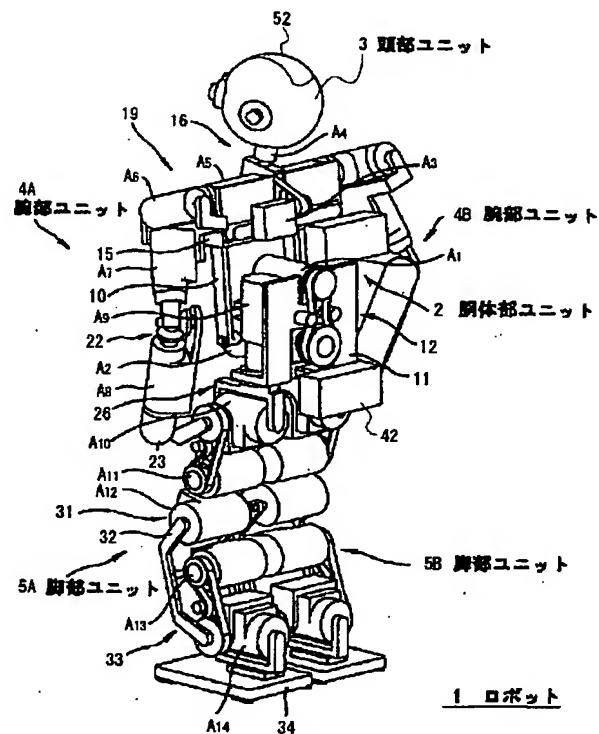


図2 本実施の形態によるロボットの外觀構成 (2)

【図6】

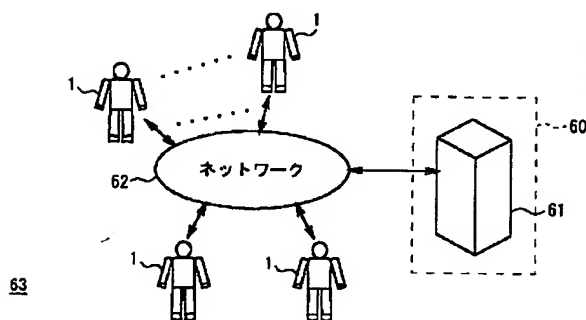


図6 本実施の形態による対話制御システムの構成

【図7】

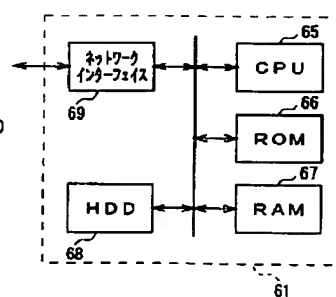


図7 コンテンツサーバの構成

【図9】

SID	名前
2	フジタ
5	ヨシダ

図9 SIDと名前との関連付け

【图 4】

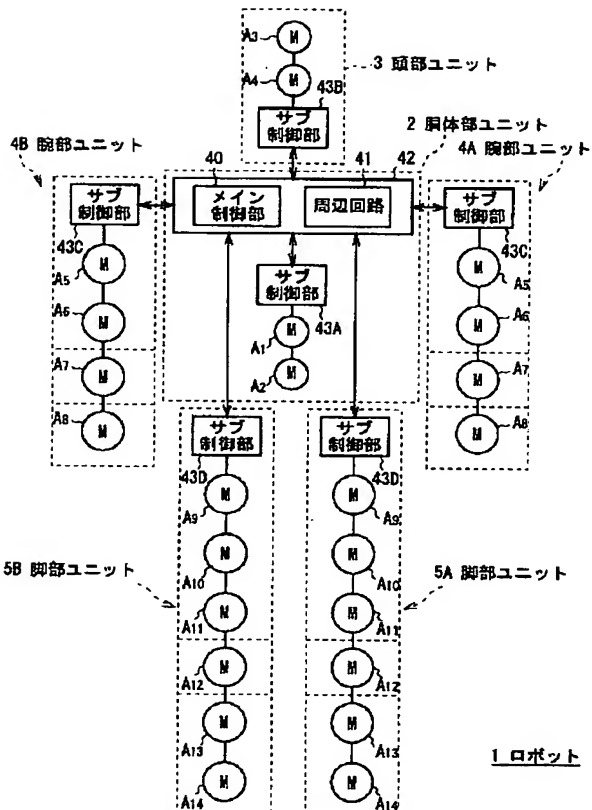


図4 ロボットの内部構成(1)

【圖 10】

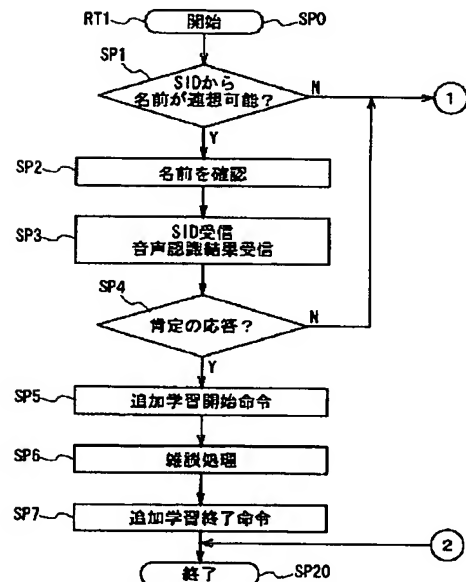


圖 10 名前學習處理手順 (1)

【图 1-1】

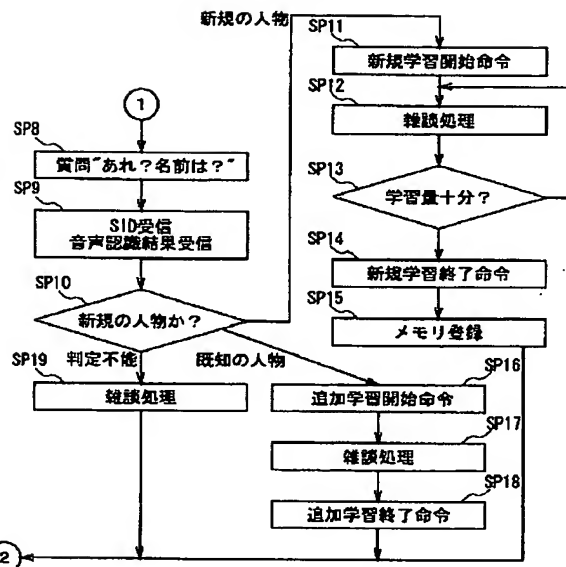


图 1-1 名前学習処理手順 (2)

【例 13】

R: フジさんですね。 R: あ、フジさんですね。
H: はい。そうです。 H: はい。そうです。
R: 今日はいい天気ですね。 R: あれ、名前を教えてください。
H: そうですね。 H: ヤマモトさん、
R: 私はロボットです。よろしくお願ひします。 H: ヤマモトさん、今日はいい天気ですね。
H: そうですね。 H: そうですね。

图 1 3 对应例 (2)

【图 15】

(R: あ、フジタさんですね。)
H: いいえ違います。
R: あれ、名前を教えてください。
H: ヨシダです。
R: ああヨシダさんですね。思い出しましたよ。今日はいい天気ですね。
H: そうですね。
R: 前回はえーと、いつ会いましたっけ？
H: ええと、おとといじゃなかったかな？

图 15 对应例 (3)

【图 19】

```

$col1 = [kono | sono] iro wa:
$this = kore [ga | wa | mo]:
$neg = (Chagau | ie) $sil:
$null = $sil:
$des = (desu | da) [yo] | yo:
$not = janai [yo]:
$color1 = $null | $neg | [$neg] $col1 | [$neg] $this:
$color2 = [iro] (desu | janai | da) [yo]:
$pat1 = $color $arabazu $color2:

```

圖 19 文法規則

【图 16】

S I D	名前
2	フジタ
5	ヨシダ
6	ヤマモト

R: あ、フジタさんですね。
H: いいえ違います。
R: あれ、名前を教えてください。
H: ヤマモトです。
R: ああそうですか。元気ですか？
H: ええ元気ですよ。

図 14 SIDと名前の新規登録

图 16 对应例 (4)

【圖 18】

見出し	音価系列
boku [僕]	boku
chigau [違う]	chigau
doko [どこ]	doko
genki [元気]	genki
iro [色]	iro
janai [じゃない]	janai
kirai [嫌い]	kirai
kudasai [ください]	kudasai

图 18 单語辞書

【図17】

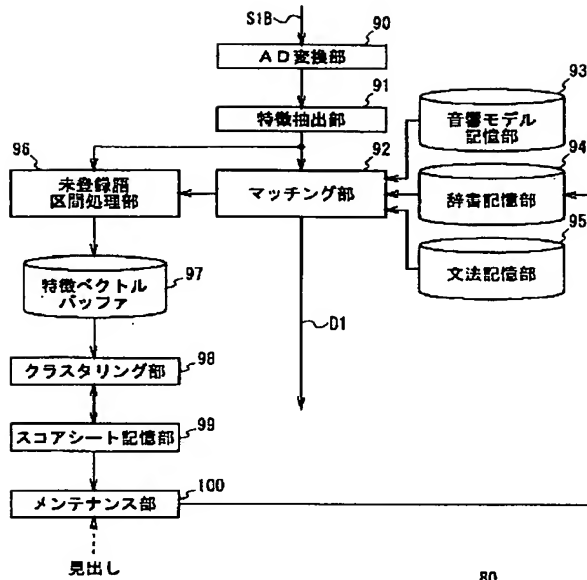


図17 音声認識部の構成

【図21】

ID	音韻 系列	クラス ナンバ	代表 メンバID	スコア(距離)					
				1	2	3	...	N	N+1
1	...	1	1	s(1,1)	s(1,2)	s(1,3)	...	s(1,N)	s(1,N+1)
2	...	2	2	s(2,1)	s(2,2)	s(2,3)	...	s(2,N)	s(2,N+1)
3	...	1	1	s(3,1)	s(3,2)	s(3,3)	...	s(3,N)	s(3,N+1)
.
.
.
N	...	1	1	s(N,1)	s(N,2)	s(N,3)	...	s(N,N)	s(N,N+1)
N+1	...	2	2	s(N+1,1)	s(N+1,2)	s(N+1,3)	...	s(N+1,N)	s(N+1,N+1)

図21 スコアシート

【図27】

ユーザ名		レベル	すでに遊んだID	遊んだ回数
○山△子	なぞなぞ	2	1, 3, ...	10
	山手線	4	1, 2, ...	5
□山×男	なぞなぞ	5	3, 4, ...	30
	山手線	2	2, 5	2

図27 プロファイルデータ

【図20】

ID	音韻系列	特徴ベクトル系列
1	---	---
2	---	---
.	.	.
.	.	.
N	---	---
N+1	---	---

図20 特徴ベクトルバッファ

【図22】

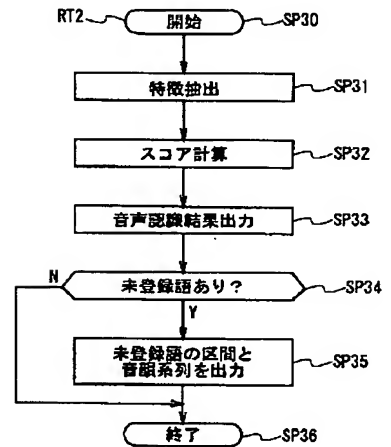


図22 音声認識処理手順

【図25】

音韻系列	発話単語
doroa:	風呂×1:
kuro	風呂×3:
bfuro	風呂×20:
hohn	本×18:
hohn	本×6:
hohns	本×10:
hohns: su	本×4:
ohh	オレンジ×1:本×19:
hohgdawasoNre: a:	本×2:
a:modori:	緑色×11:
onidori:	緑色×10:
e:imidori:	緑色×3:
imidori:	緑色×5:
a:imidori:iroiresu	緑色×4:
Nro:ka	廊下×10:
Nro:kaNa	廊下×10:

図25 シミュレーション結果

【図23】

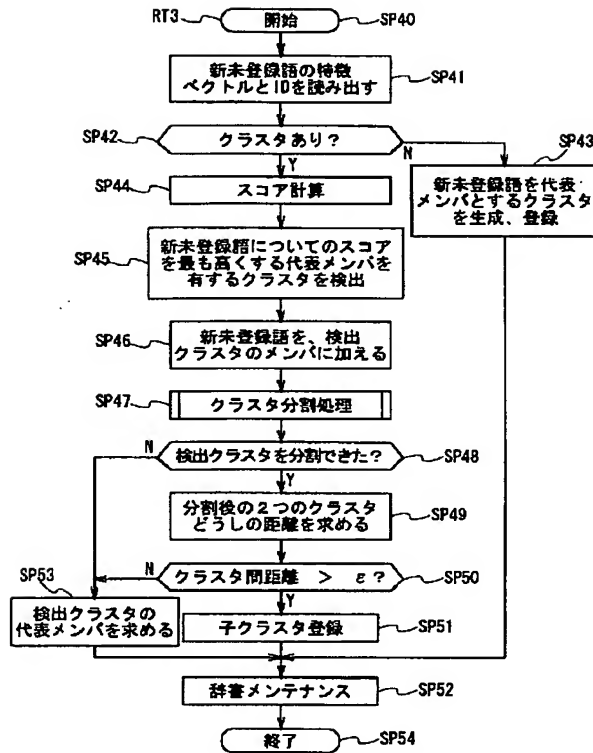


図23 未登録語処理手順

【図28】

コンテンツデータ				オプションデータ	
ID	問題	答え	理由	難易度	人気
1	4歳と5歳の子供しか住んでいない外国の都市は?	シカゴ	4歳と5歳でシカゴだよ	2	5
2	少ししか人が乗っていないのに一杯な車はなんだ?	救急車	キューキューで一杯だよ	3	7
3	家の中で暖房が効かない場所はどこだ?	玄関	寒い寒さで寒いだよ	5	3
4	落ち込んでいても2回食べると元気になるのは?	のり	2回でのりのりだよ	2	8

図28 コンテンツデータ

【図24】

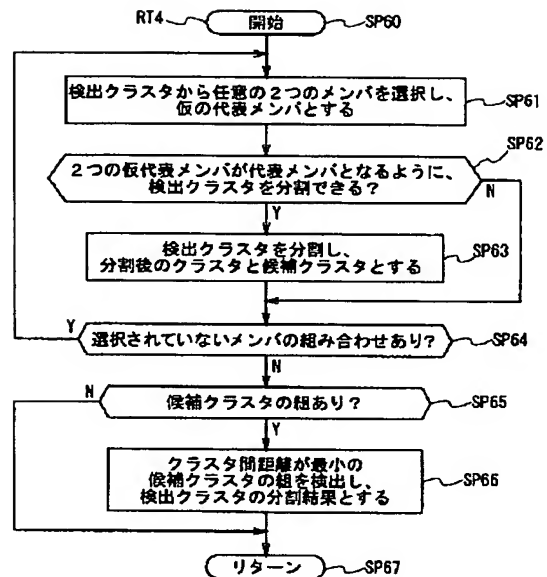


図24 クラスタ分割処理手順

【図26】

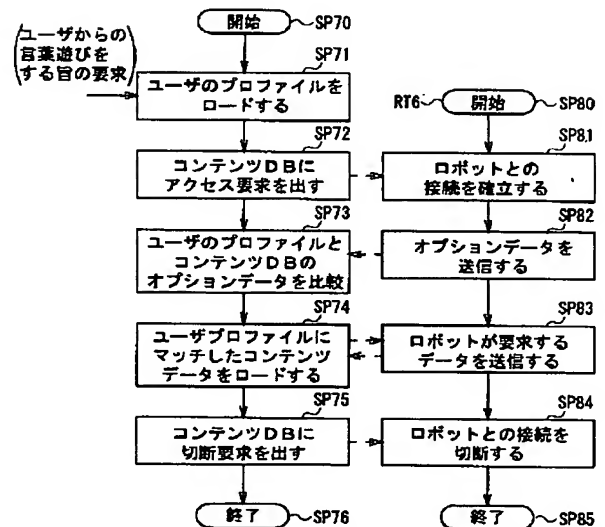


図26 コンテンツデータ取得処理手順及びコンテンツデータ提供処理手順

【図29】

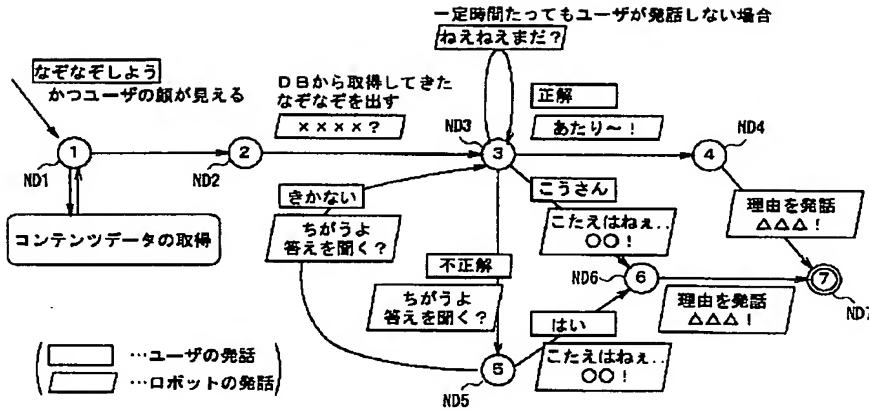


図29 言葉遊びによる対話シーケンス(1)

【図30】

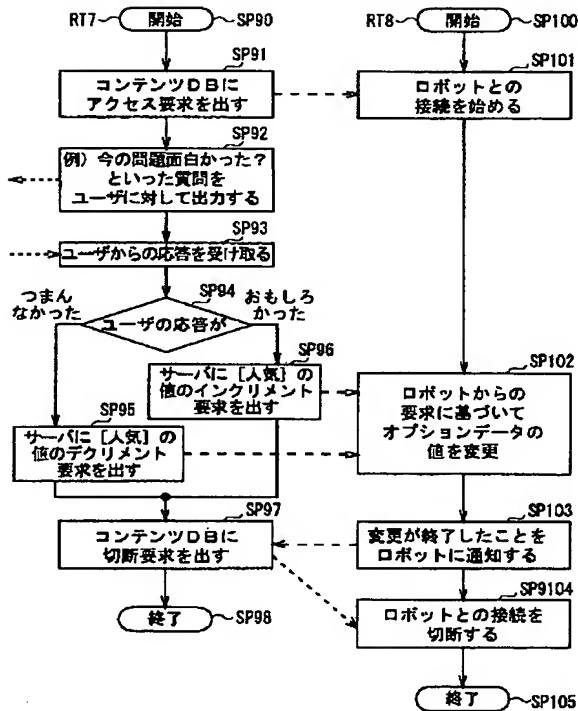


図30 人気指標集計処理手順及びオプションデータ更新処理手順

【図31】

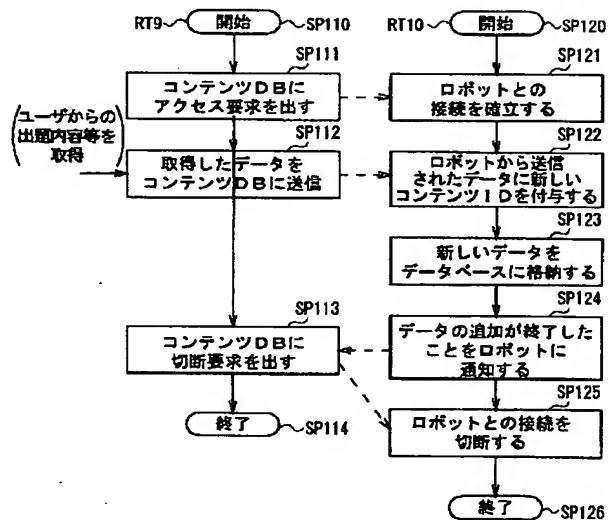


図31 コンテンツ収集処理手順及びコンテンツデータ追加登録処理手順

【図32】

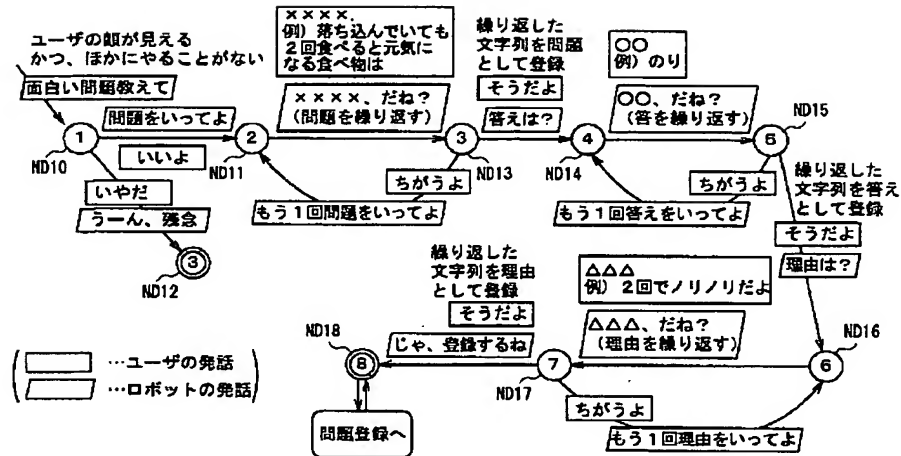


図32 言葉遊びによる対話シーケンス(2)

フロントページの続き

(51) Int. Cl. 7

識別記号

F I

テマコード(参考)

G 1 0 L 15/00
15/06
15/20
17/00

G 1 0 L 3/00

5 5 1 A

R

5 3 1 P

5 2 1 J

5 4 5 A

(72) 発明者 山田 敬一

東京都品川区北品川6丁目7番35号ソニー
株式会社内

Fターム(参考) 2C150 BA11 CA01 CA02 DA04 DA05

DA24 DA26 DA27 DA28 DF03

DF04 DF06 DF33 ED10 ED42

ED47 ED52 EF03 EF07 EF09

EF13 EF16 EF17 EF22 EF23

EF28 EF29 EF33 EF36

3C007 AS36 CS08 JS03 KS39 MT14

WA03 WA13 WB19 WC01 WC03

WC07

5D015 KK02 KK04 LL02

5D045 AB11